VOL. 16 No. 11, pp. 309-340 MAY 1955

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

# REFLECTION AND IMPEDANCE MEASUREMENTS BY MEANS OF A LONG TRANSMISSION LINE

by J. C. van den HOOGENBAND and J. STOLK.

621.317.332.6

*In systems employing metre (and shorter) waves, it is often necessary to match, say, an aerial to a transmission line, or a transmission line to a receiver. A method providing a direct visual indication of matching, by means of an oscilloscope, is described in this article. This method also enables the reflection factor to be deduced quite readily from the oscillograms, and hence the correct line-terminating impedance. Given this impedance, it is a relatively simple matter to determine the measures necessary to ensure correct matching.*

Many methods of measuring impedances at very high frequencies are known. One of them, i.e. measurement of the unknown impedance in a bridge circuit, is used in several different forms. However, these measurements present some difficulty at frequencies higher than about 300 Mc/s (decimetre waves), owing to the fact that here the wavelength is of the same order of magnitude as the dimensions of the object on test and the length of the connecting cable between it and the measuring equipment. The effects which therefore occur at such frequencies may be attributed to the stray self-inductance and capacitance; corrections to measurements, although possible, are never very accurate.

According to another principle, the impedance can be measured from the detuning and damping effects when it is in parallel with a tuned circuit such as a lecher system (transmission line)[1]. Again, however, this method is accurate only when the impedance to be measured is either appreciably lower, or appreciably higher, than the characteristic impedance of the transmission line.

Another method, in fact an elaboration of the one just described, produces accurate results in the case of an unknown impedance of the same order of magnitude as the characteristic impedance $\zeta$ (i.e. from about 0.1 $\zeta$ to 10 $\zeta$). Here, the voltage distribution along an untuned transmission line is determined by means of a movable detector (standing wave indicator.) From the ratio of the voltage maxima and minima (standing wave ratio), and the position of the minima, the reflection coefficient (to be defined later) can be deduced. Both components of the required impedance can then be calculated from this coefficient [2] [3]. In many cases the object of the measurement is to determine the reflection factor itself; it indicates in how far the transmission line is terminated with an impedance equivalent to its characteristic impedance, that is, in how far these impedances are matched.

In very high frequency applications, many of the objects tested are specially designed for connection to a transmission line [4]. If the characteristic impedance of the test-object is the same as that of the transmission line, the result of the measurement will be a true picture of the behaviour of the line and the impedance combined, which is precisely the information required in practice.

---

[1] J. M. van Hofweegen, The measurement of impedances, particularly on decimeter waves, Philips tech. Rev. **8**, 16-24, 1946.

[2] J. M. van Hofweegen, Impedance measurements with a non-tuned Lecher system. Philips tech. Rev. **8**, 278-286, 1946.

[3] H. J. Lindenhovius, The measurement of impedances at high frequencies and applications of the standing-wave indicator, T. Ned. Radiogenootschap **12**, 65-82, 1947 (in Dutch).

[4] A. E. Pannenborg, A measuring arrangement for wave-guides, Philips tech. Rev. **12**, 15-24, 1950/51.

Although measurements carried out with the standing-wave indicator produce quite accurate results, the mechanical quality of the indicator itself must conform to very stringent requirements. The method that will now be described, although perhaps less accurate, is much simpler both as regards the equipment required and the actual measurement. It resembles the method described in article [2]) in as far as the impedance to be determined is deduced from the reflection coefficient. However, the new method offers at least one considerable advantage, that is, that the reflection coefficient itself is readily determined from oscillograms; the oscillogram shows whether matching has been effected or not. Hence this method is most suitable in cases where the matching of a transmission line to, say, a signal generator or an aerial is to be adjusted or checked quickly over a wide range of frequencies.

The principle employed here for determining the absolute value of the reflection coefficient has already been described by Bayer [5]). However, he did not suggest the possibility of calculating the unknown impedance entirely from the reflection coefficient. The object of the present article is to develop the method outlined by Bauer, and to formulate the results mathematically and describe one or two practical applications.

### Incident and reflected waves in a transmission line

Using the method of complex quantities, the voltage and current at an arbitrary point ($y$) on a transmission line (*fig. 1*) may be defined as $V(y)e^{j\omega t}$ and $I(y)e^{j\omega t}$, respectively, where $\omega$ is the angular frequency and $V(y)$ and $I(y)$ are complex quantities depending solely on the position ($y$), that is, independent of the time ($t$).

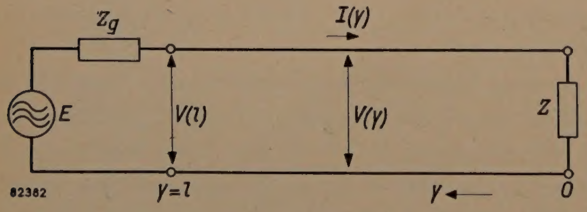


Fig. 1. Transmission line fed by a generator of e.m.f. $E$ and internal impedance $Z_g$, and terminated in an impedance $Z$.

For $V(y)$ and $I(y)$ we have the relations

$$V(y) = A\,e^{\gamma y} + B\,e^{-\gamma y} \quad . \; . \; . \; . \; . \quad (1)$$

and

$$I(y) = \frac{A}{\zeta}\,e^{\gamma y} - \frac{B}{\zeta}\,e^{-\gamma y}, \quad . \; . \; . \quad (2)$$

where $A$ and $B$ are complex constants, independent of $t$ and $y$, the precise values of which are governed by such boundary conditions as the manner in which the system is supplied and terminated. Also, $\gamma$ is the propagation constant and $\zeta$ is the characteristic impedance of the transmission line.

In general, $\gamma$ and $\zeta$ are complex; $\gamma$ may then be written as:

$$\gamma = \alpha + j\beta, \quad . \; . \; . \; . \; . \; . \; . \quad (3)$$

where $\alpha$ is the attenuation constant (expressed in nepers per metre) and $\beta$ the phase constant (in radians per metre). If $v$ is the phase velocity of the waves in the transmission line and $f$ is the frequency, we have:

$$\beta = \frac{\omega}{v} = \frac{2\pi f}{v}. \quad . \; . \; . \; . \; . \; . \quad (4)$$

If no loss occurs in the system, $\gamma = j\beta$. The characteristic impedance $\zeta$ is then a real quantity i.e. a pure resistance; in the following it is assumed, unless otherwise stated, that there is no loss in the system (zero attenuation).

From (1) and (2), we have at the position of the unknown load impedance $Z$, i.e. at $y = 0$:

$$V(0) = A + B \quad \text{and} \quad I(0) = \frac{A-B}{\zeta}.$$

Also,

$$Z = \frac{V(0)}{I(0)}, \quad \text{hence} \quad Z = \frac{A+B}{A-B}\,\zeta,$$

or

$$B = A\,\frac{Z-\zeta}{Z+\zeta} = Ar,$$

where

$$r = \frac{Z-\zeta}{Z+\zeta} = |r|\,e^{j\varphi} \quad . \; . \; . \; . \; . \quad (5)$$

is complex.

Substituting $B = Ar$ in (1) and (2), we have:

$$V(y) = A(e^{\gamma y} + r\,e^{-\gamma y}) \quad . \; . \; . \; . \quad (6a)$$

and

$$I(y) = \frac{A}{\zeta}\,(e^{\gamma y} - r\,e^{-\gamma y}). \quad . \; . \; . \quad (6b)$$

It is seen from these equations that the voltage distribution in the line may be considered as the superposition of an outgoing wave $Ae^{\gamma y}$ and a returning wave $Are^{-\gamma y}$; the latter may be imagined to arise from a partial reflection of the outgoing wave by the load impedance $Z$. Since $r$ is a measure of this reflection, it is termed the *reflection coefficient*. From (5), $r$ is determined uniquely by the ratio of the terminating impedance $Z$ to the characteristic impedance $\zeta$. Given $Z = \zeta$, then, $r = 0$. In this case, we have only an outgoing wave ("travelling wave"). If $Z = 0$ (end short-circuited), $r = -1$; if $Z = \infty$ (open-circuit end), $r = +1$. In both cases the reflected wave has the same amplitude as the outgoing wave, i.e. there is *total reflection*. Thus a "standing wave" is produced.

---

[5]) J. A. Bauer, Special applications of ultra-high frequency wide-band sweep generators, R. C. A. Rev. 8, 564-575, 1947.

### *Reasons for the undesirability of reflection*

Provided that the condition that the impedance $Z_g$ of the generator (fig. 1) shall equal $\zeta$ is satisfied, the power carried by the outgoing wave will be $(\frac{1}{2}E)^2/\zeta$, regardless of the value of the load impedance $Z$. From this "available power" $P_1 = (\frac{1}{2}E)^2/\zeta$, a part $P_2 = (\frac{1}{2}|r|E)^2/\zeta$ is reflected by $Z$ and carried back by the reflected wave. Accordingly, the ratio:

$$\eta = \frac{P_1 - P_2}{P_1} = 1 - |r|^2$$

is the efficiency of the energy transfer.

The square root of $\eta$ is equal to the ratio of the voltage across the resistive part ($R$) of the load impedance $Z$ to the voltage that would be obtained across the same resistance $R$ with proper matching, that is, if all the energy were concentrated in $R$. Hence $\sqrt{\eta}$ may be termed the "efficiency" of the voltage transfer. It will be seen from *fig. 2*, which shows $\eta$ and $\sqrt{\eta}$ plotted against $|r|$, that in general the reflections should be minimized by a proper match.

There are even stronger reasons to minimize reflections. In the case of a television receiver, for example, the cable between set and aerial must be matched at *both* ends; otherwise, reflected waves, as well as the desired "principal wave", will enter the receiver. These reflected waves would distort abrupt black to white transients in the picture.

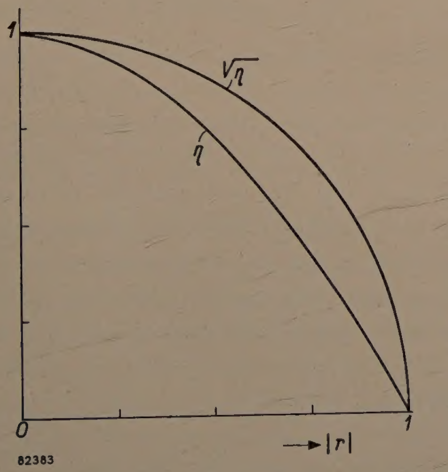


Fig. 2. Efficiency $\eta$ of the energy transfer, and "voltage efficiency" $\sqrt{\eta}$, of a transmission line, plotted against the modulus $|r|$ of the reflection coefficient. High values of $\eta$ and $\sqrt{\eta}$ are possible only when $|r|$ is small.

### *Smith's chart*

Equation (5) can be appreciably clarified with the help of a Smith chart [6] (*fig. 3*). Here, $r$ is plotted in the complex plane. Given $Z = R + jX$, i.e. $Z/\zeta = R/\zeta + jX/\zeta$, the locus described by $r$ when $X/\zeta$ varies and $R/\zeta$ remains constant can be determined with the aid of equation (5). For each value of $R/\zeta$ a circle is obtained whose centre lies on the real axis and which passes through the point $+1$. We thus get a system of circles all passing through the point $+1$. If $X/\zeta$ is taken as the parameter and $R/\zeta$ as the variable, another system of circles orthogonal to the first system will be obtained.

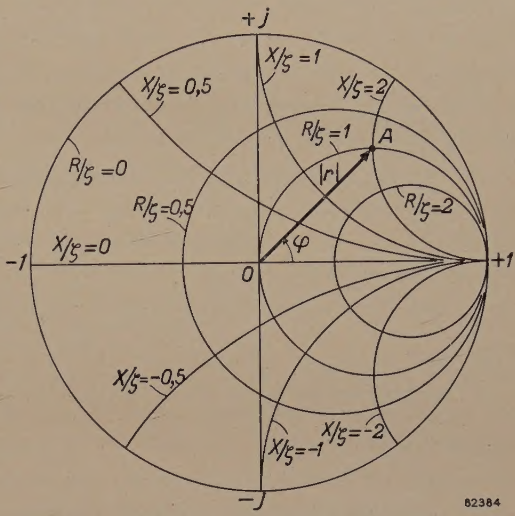


Fig. 3. Smith chart, showing the two orthogonal families of circles, the one referring to constant $R/\zeta$ and the other to constant $X/\zeta$. $A$ is the point of intersection of the circles corresponding to $R/\zeta = 1$ and $X/\zeta = 2$ ($R$ being the resistance, and $X$ the reactance of the terminating impedance $Z$ from fig. 1). Vector $OA$ is the reflection coefficient $r$ with modulus $|r|$ and argument $\varphi$.

Now, for a given value of $\zeta$, the point associated with any impedance $Z = R + jX$, at which the circle corresponding to $R/\zeta$ cuts that corresponding to $X/\zeta$, can be established. This point determines the modulus $|r|$ and the argument $\varphi$ of the reflection factor $r$.

Since $|r| \leq 1$, the entire diagram is contained within a circle of radius unity. The dimensionless parameters $R/\zeta$ and $X/\zeta$ enable the diagram to be adapted to any characteristic impedance $\zeta$; the reading becomes inaccurate only at very low or very high values of these parameters.

---

[6] P. H. Smith, An improved transmission-line calculator, Electronics **17**, 130-133 and 318-325, Jan. 1944. A similar diagram was worked out independently during the war by J. M. van Hofweegen, who describes it in his article [2]), pages 283-284.

Many results derived in the theory of transmission lines can be illustrated by means of the Smith chart. For example, the impedance of a line terminated in a given impedance $Z = R + \mathrm{j}X$ can be determined as follows (*fig. 4a*). If this line be cut at a certain point $y = y_1$ and there terminated with an impedance matching the input impedance $Z_1$ of the piece cut off (fig. 4*b*), the outgoing and reflected waves in the remainder of the line will not change. Hence the amplitudes the outgoing and reflected waves at point $y = y_1$ will be the same after as before the line is cut, viz. (from (6a)) $A\mathrm{e}^{\gamma y_1}$ and $Ar\mathrm{e}^{-\gamma y_1}$ respectively. In the case to which fig. 4*b* refers, then, a reflection coefficient of

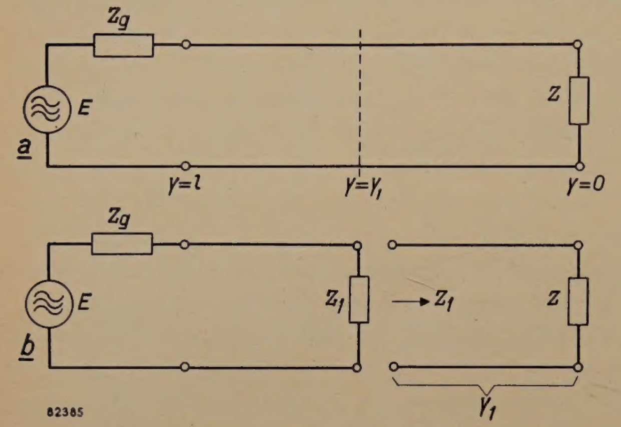


Fig. 4. *a*) Transmission line as shown in fig. 1. *b*) A certain length $y_1$ of the system is cut off and replaced by the equivalent impedance $(Z_1)$.

$$r_1 = \frac{A\,r\,\mathrm{e}^{-\gamma y_1}}{A\,\mathrm{e}^{\gamma y_1}} = r\,\mathrm{e}^{-2\gamma y_1} \quad . \; . \; . \; . \; . \quad (7)$$

is required at point $y = y_1$. From (3) and (4), assuming the line to be without loss, we may write for formula (7):

$$r_1 = r\,\mathrm{e}^{-\mathrm{j}4\pi f y_1/v} \quad . \; . \; . \; . \; . \; . \quad (8)$$

Accordingly, $r_1$ is determined by rotating the vector $r$ in the Smith chart through an angle $4\pi f y_1/v$ (*fig. 5*). The real and imaginary parts of impedance $Z_1$, seen at the input of the cut-off portion of the line, are then read from the families of circles associated with constant $R/\zeta$ and constant $X/\zeta$.

In this way we find that, say, at $\zeta = 100$ ohms and $Z = 100 + \mathrm{j} \times 100$ ohms (that is, $R/\zeta$ and $X/\zeta$ both unity, point $A$), the modulus of the reflection factor $r$ is $|r| = 0.447$ and its argument is $\varphi = 63°$. If the line has a length $l$ equal to $^1/_4$ wavelength ("$\frac{1}{4}\lambda$-line"), the angle of rotation $4\pi f\,l/v$ is $\pi$. The reading from fig. 5 corresponding to this value (point $C$) is: $R/\zeta = \frac{1}{2}$ and $X/\zeta = -\frac{1}{2}$; therefore $Z_1 = 50 - \mathrm{j} \times 50$ ohms.

A similar method is employed to determine the impedance $Z_l$ of a cable terminated with impedance $Z$, considered as a load on the voltage source at the input. Let us assume that this source is a generator of e.m.f. $E$ and internal impedance $Z_\mathrm{g}$ (fig. 1). The vector $r$ in the diagram must then be rotated through an angle $4\pi f\,l/v$. If $Z_\mathrm{g} = \zeta$, the voltage $V(l)$ on the input terminals of the cable is:

$$V(l) = \frac{Z_l}{Z_l + \zeta}\,E = \frac{\frac{1}{2}Z_l + \frac{1}{2}\zeta + \frac{1}{2}Z_l - \frac{1}{2}\zeta}{Z_l + \zeta}\,E$$
$$= \tfrac{1}{2}(1 + r_l)E, \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad (9)$$

where, from (5), $r_l = (Z_l - \zeta)/(Z_l + \zeta)$ is the reflection factor at the input of the cable (where $y = l$). For $r_l$, from (8) we may write

$$r_l = r\,\mathrm{e}^{-\mathrm{j}4\pi f l/v},$$

therefore

$$V(l) = \tfrac{1}{2}E\,(1 + |r|\,\mathrm{e}^{\mathrm{j}\varphi} \cdot \mathrm{e}^{-4\pi f l/v}).$$

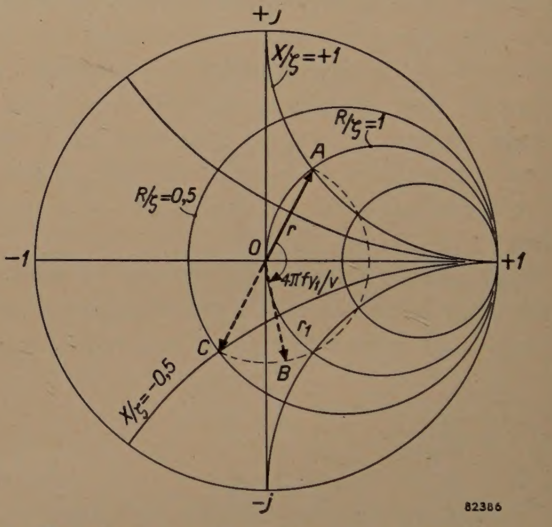


Fig. 5. The reflection coefficient $r_1 = OB$ of the remainder of the transmission line, shown in fig. 4*b*, is determined by rotating vector $r = OA$ in the Smith chart through an angle $4\pi f y_1/v$.

It is seen from (9) that vector $PQ$, representing voltage $V(l)$ in the diagram (see *fig. 6a*), arises from the summation of constant vector $PO$ ($= \frac{1}{2}\,E$) and vector $OQ$ ($= \frac{1}{2}\,r_l\,E$). With a suitable choice of scale, $\frac{1}{2}\,E$ can be made equal to the radius of the diagram.

Any increase in the frequency $f$ will be indicated by a rotation of $Q$ around the circle with radius $\frac{1}{2}|r_l|E$, from which the dependence of the input voltage $|V(l)|$ on the frequency can be deduced (fig. 6*b*).

It will be seen from the above that the amplitude of the outgoing wave ($\frac{1}{2}E$) does not depend upon the reflection taking place at the end. However, it should be noted that this does not hold good when $Z_g \neq \zeta$, since the wave reflected at the end is then reflected a second time by the voltage source, and so contributes to the outgoing wave. In this case the calculations are complex, and are in fact entirely beyond the scope of this article. However, one or two results will be referred to later.

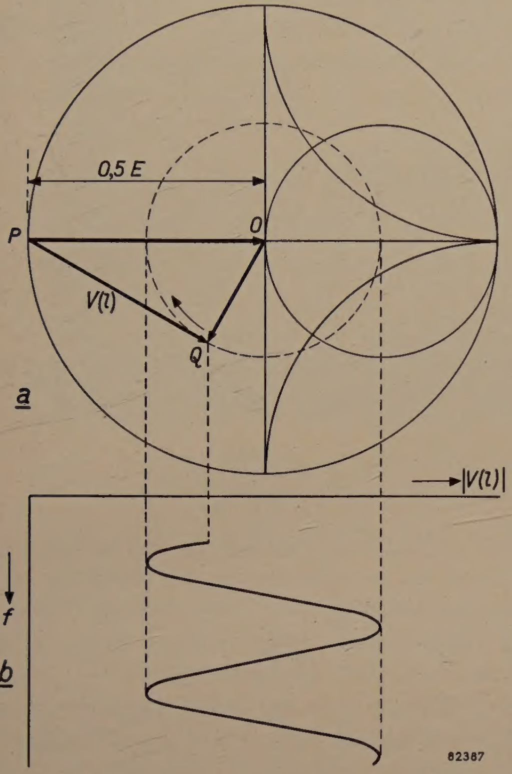


Fig. 6. *a*) $V(l) = \frac{1}{2}(1 + r_l)E$ as demonstrated in the Smith chart. When the frequency increases, $Q$ moves in the clockwise direction around the dotted circle with radius $\frac{1}{2}|r_l|E$. *b*) The modulus of $V(l)$, corresponding to line $PQ$ in (*a*), plotted against the frequency $f$.

## The measurement of reflections

The method described in previous articles [2] [3]) involved a fixed frequency and a movable detector. Here, a variable frequency and a stationary detector are employed.

The modified method now to be described is specially designed to facilitate matching over a wide range of frequencies. Other variants of the method more suitable where accuracy is preferred to ready interpretation and speed, will be discussed at the end of the this article.

A block diagram of the measuring equipment is shown in *fig. 7*.

Here, a transmission line whose characteristic impedance is known, is connected, with a detector in parallel, to a "wobbulator", that is, a signal generator whose frequency varies periodically, in this case with the mains frequency. The transmission line is either a coaxial or an unshielded twin cable, many times longer than the mean wavelength. It is terminated with the unknown impedance $Z$. The equipment also includes an oscilloscope. The time-base voltage of this instrument is synchronized to the frequency modulation of the wobbulator, so that the abscissa of the oscillogram is the frequency co-ordinate. The detector produces a D.C. voltage fluctuating at low frequency; the A.C. component of this voltage, after amplification, produces the vertical deflection. A diagram to show the type of oscillogram obtained with this system is given in fig. 6*b*.

With a circuit in which the signal to produce the vertical deflection is present only during the forward trace, the horizontal axis of the oscillogram will be described during the return trace.

The oscillograms obtained in one or two simple cases will now be described.

*a*) The impedance $Z$ is a pure resistance $R$. Then:

$$r = \frac{R-\zeta}{R+\zeta},$$

Hence $\varphi = 0$ when $R > \zeta$, and $\varphi = \pi$ when $R < \zeta$. Accordingly, the head of vector $r$ is always on the real axis, at $-1$ when $R = 0$, at $-\frac{1}{2}$ when $R = \frac{1}{3}\zeta$, at 0 when $R = \zeta$, at $+\frac{1}{2}$ when $R = 3\zeta$ and at $+1$ when $R = \infty$ (see *fig. 8a*).

When the frequency varies, point $Q$ in diagram 6*a* describes a circle about 0, the radius of this circle being governed by $R$. The length of line $PQ$ is then the modulus of vector $V(l)$. The oscillogram of $|V(l)|$ as a function of $(4\pi l/v)f$ for the above-mentioned values of $R$ will then be as shown in fig. 8*b*.

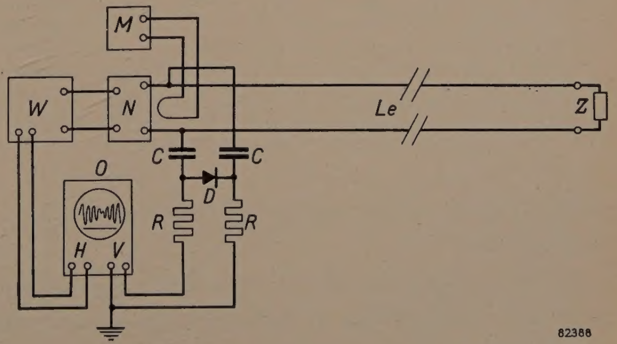


Fig. 7. Circuit employed to determine the reflection coefficient of an unknown impedance ($Z$) terminating the line *Le* ($\zeta$ = 300 ohms). *W* wobbulator, *N* matching network, *V* detector, *O* oscilloscope (horizontal deflection synchronized with the frequency variation of the wobbulator, vertical deflection proportional to the A.C. output voltage of the detector), *M* marker (described later), *C* isolating capacitors (100 pF), *R* resistors (10,000 ohms), to prevent H.F. voltage from entering the oscilloscope.

*b*) The impedance $Z$ is pure a reactance j$X$. Then:

$$r = \frac{\mathrm{j}X - \zeta}{\mathrm{j}X + \zeta}.$$

Hence

$$|r| = 1, \qquad \varphi = \pi - 2\tan^{-1}\frac{X}{\zeta}.$$

The fact that $|r| = 1$, shows that in the case of a purely reactive terminating impedance the head of vector $r$ is always on the periphery of the Smith chart (fig. 8*c*). Fig. 8*d* represents the oscillogram of $|V(l)|$ as a function of $(4\pi l/v)f$ for three values of $X$. Since only the argument ($\varphi$) of $r$ varies with $X$, a change of $X$ merely produces a horizontal displacement of the oscillogram.

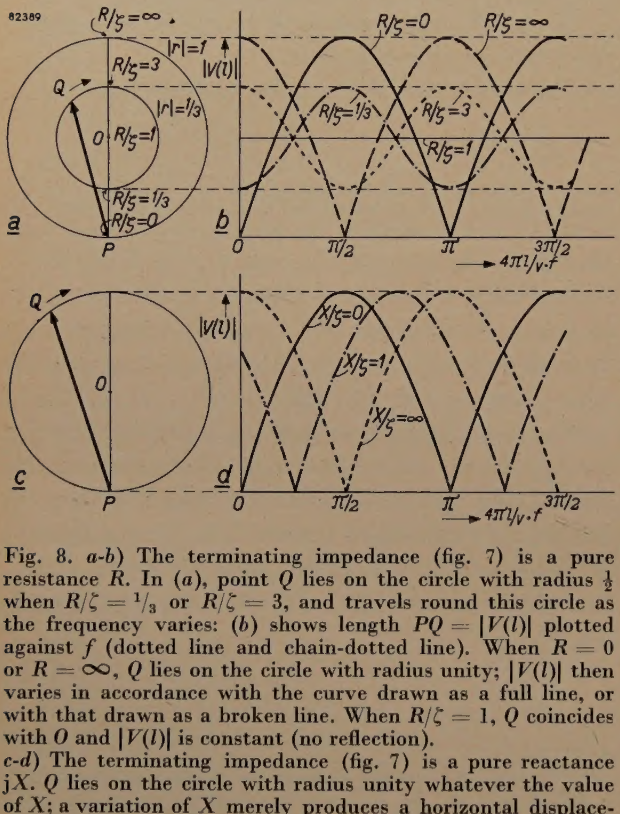


Fig. 8. *a-b*) The terminating impedance (fig. 7) is a pure resistance $R$. In (*a*), point $Q$ lies on the circle with radius $\frac{1}{2}$ when $R/\zeta = 1/3$ or $R/\zeta = 3$, and travels round this circle as the frequency varies: (*b*) shows length $PQ = |V(l)|$ plotted against $f$ (dotted line and chain-dotted line). When $R = 0$ or $R = \infty$, $Q$ lies on the circle with radius unity; $|V(l)|$ then varies in accordance with the curve drawn as a full line, or with that drawn as a broken line. When $R/\zeta = 1$, $Q$ coincides with $O$ and $|V(l)|$ is constant (no reflection).
*c-d*) The terminating impedance (fig. 7) is a pure reactance j$X$. $Q$ lies on the circle with radius unity whatever the value of $X$; a variation of $X$ merely produces a horizontal displacement of $|V(l)|$ as a function of $f$.

### *Determination of the reflection coefficient from the oscillograms*

The determination of the reflection coefficient $r$ from the oscillograms is based on the following considerations:

*a*) From (9), $|V(l)|$ fluctuates between $\frac{1}{2}E(1 + |r|)$ and $\frac{1}{2}E(1 - |r|)$. Accordingly, the oscillogram shows a ripple whose amplitude $a_r$ (peak to peak, *fig. 9a*) is proportional to the variation of $|V(l)|$; hence:

$$a_r = g \cdot \tfrac{1}{2}E\{(1 + |r|) - (1 - |r|)\} = gE|r|.$$

The proportionality factor ($g$) is determined by introducing a known reflection coefficient, e.g. total reflection ($|r| = 1$, fig. 9*b*), say, by leaving the end of the line open. The ripple amplitude ($a_0$) is then $gE$, so that:

$$|r| = \frac{a_r}{a_0}. \quad \ldots \ldots \ldots \quad (10)$$

Accordingly, the ratio of the two ripple amplitudes is equal to the modulus of the reflection coefficient
*b*) The oscillogram is periodic. In the event of an increase $(\Delta f)_0$ in $f$, such that $4\pi(\Delta f)_0 l/v = 2\pi$, so that

$$(\Delta f)_0 = \frac{v}{2l}, \quad \ldots \ldots \quad (11)$$

$|V(l)|$ will return to the original value. Given a long transmission line, then, the frequency variation will be small.

*c*) The interval between the minima in fig. 8*a* and *b* is $(\Delta f)_r = (\varphi/2\pi)(\Delta f)_0$. Hence

$$\varphi = \frac{(\Delta f)_r}{(\Delta f)_0} 2\pi. \quad \ldots \ldots \quad (12)$$

Total reflection can be obtained not only by open-circuiting the end of the line, but also by short-circuiting it. In the latter case, $\pi$ should be added to the result of (12).

It is seen, then, that the ratio of the two differences in frequency to be read from the oscillograms, corresponds to the argument of the reflection coefficient. Hence the reflection coefficient itself can be fully established by means of (10) and (12) combined.

The derivation of the unknown impedance, $Z$, from the Smith chart with the aid of $|r|$ and $\varphi$ has already been described.

However, the method that will now be considered is more accurate than the mere reading of $(\Delta f)_r$ and $(\Delta f)_0$ from oscillograms. An auxiliary oscillator

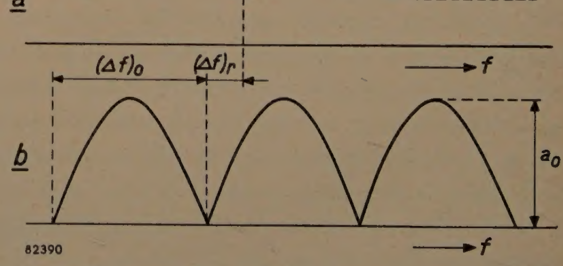


Fig. 9. Oscillograms obtained with the circuit shown in fig. 7, (*a*) employing an arbitrary terminating impedance. (*b*) with the end of the test line open. The ratio of the ripple amplitudes ($a_r$ and $a_0$) is the modulus of the reflection coefficient, and $2\pi$ times the ratio of the differences in frequency $(\Delta f)_r$ and $(\Delta f)_0$ is the argument of this coefficient ($\varphi$).

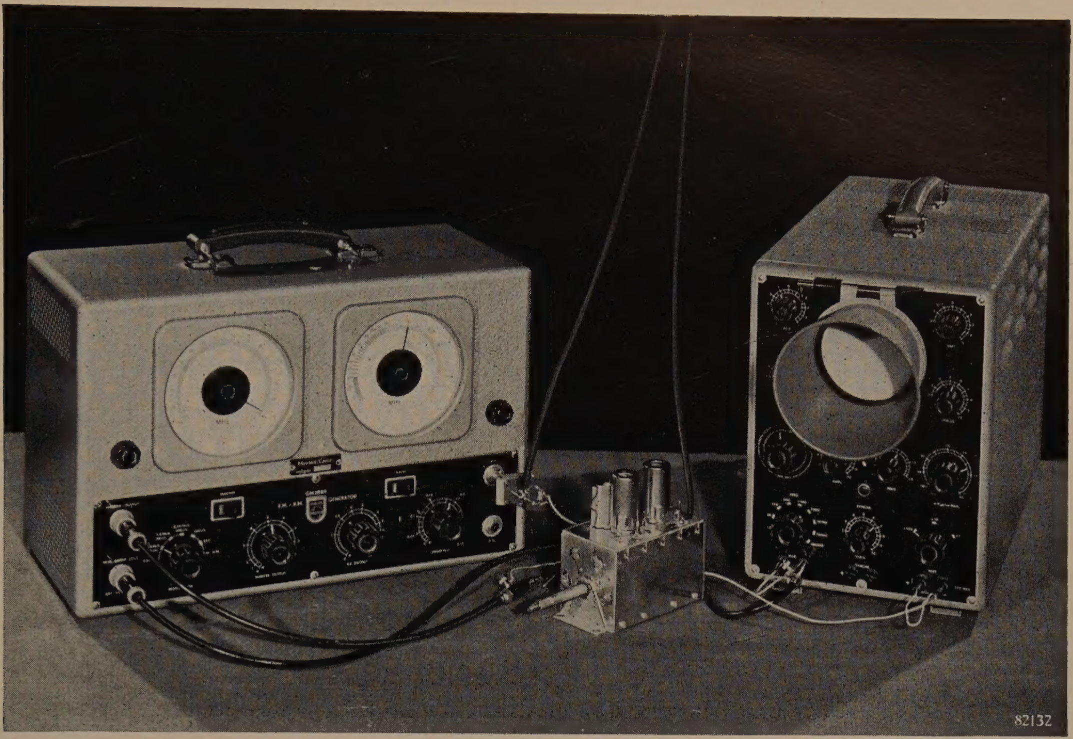


Fig. 10. Equipment for the measurement of reflections. From left to right: the wobbulator (with built-in marker), type GM 2889, the object under test (here a channel-selector for a TV receiver) and the cathode-ray oscilloscope. Note the input and output ends of the 60 metre unshielded twin lead employed as the transmission line.

($M$ in fig. 7), with facilities for varying and reading the frequency accurately, is coupled loosely to the measuring system. This oscillator is then tuned to a frequency within the frequency variation of the wobbulator. Each time the wobbulator frequency passes that of the auxiliary oscillator, a beat occurs producing a small irregularity in the line in the oscillogram (hence the auxiliary oscillator is termed a "marker"). The marker enables $(\Delta f)_r$ and $(\Delta f)_0$ to be determined very accurately as differences between consecutive adjustments of the marker.

The measuring equipment is shown in *fig. 10.*

### *Effect of attenuation*

The impedance to be measured, $Z$, is generally dependent on the frequency; hence $|r|$ and $\varphi$ both vary with the frequency. Employing the method described here, it is necessary to vary the frequency over a range comprising several intervals $(\Delta f)_0$, and for the sake of accuracy, $|r|$ must be prevented from varying unduly within any one of these intervals. This condition is easier to satisfy the smaller the interval. From (11), however, it is necessary to increase the length ($l$) of the transmission line according as $(\Delta f)_0$ decreases, and the attenuation, so far ignored, increases with the length of the line.

We must therefore now consider the effect of attenuation on the measurements.

If the attenuation is $a$ nepers per metre, the amplitude of the reflected wave on reaching the detector will be a factor of $e^{-2al}$ lower than it would would be without attenuation. Owing to this relatively lower amplitude, the voltage $|V(l)|$ will invariably remain above zero, even in the case of total reflection (compare *fig. 11a,* relating to $a = 0$, with fig. 11$b$, which refers to $a > 0$). Accordingly, it is necessary to multiply $|r|E$ and $E$, in the expressions of $a_r$ and $a_0$, respectively, by $e^{-2al}$; at the same time, the ratio $a_r/a_0$, which, from

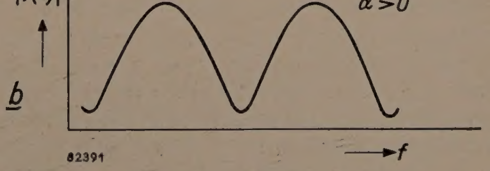


Fig. 11. $|V(l)|$, plotted against $f$ in the case of total reflection $a$) with linear deflection, but without attenuation, $b$) with linear deflection and attenuation. In case ($b$), $|V(l)|$ remains above the zero line.

(10), is equal to $|r|$, remains unchanged. In principle then, attenuation will not give rise to errors in the measurement, at least as regards $|r|$.

However, minor irregularities are invariably superimposed on the ripple produced by reflection. They may arise partly from the fact that the generator voltage varies to a certain extent with frequency, and partly from unavoidable non-uniformity in the transmission line (for example, in the case of unshielded twin lead, the supports; the effect of points of contact with the supports are minimized by suspending the cable from thin cords. The effect of insulators will be discussed later).

Such irregularities affect the result all the more, the stronger the attenuation; this imposes a limit on the effective length of the cable. A suitable length is 60 metres.

### Other effects

*Effect of mismatch between generator and transmission line*

So far it has been assumed that the internal impedance ($Z_g$) of the generator (wobbulator) matches the characteristic impedance ($\zeta$) of the measuring system. Let us now consider the variation of voltage $|V(l)|$ as a function of frequency when $Z_g$ does not match $\zeta$. The actual calculation of this variation is so complex that it is beyond the scope of this article, but one or two of the results are shown in *fig. 12*.

Fig. 12*a* depicts $|V(l)|$ plotted against $f$ for a real value of $Z_g$ (viz. $2\zeta$), and fig. 12*b* shows a similar curve for a complex value of this impedance ($Z_g = \zeta/(1-\mathrm{j})$), for $|r| = 0.33$, 0.80 and 1.0 in both cases.

When $Z_g$ is purely resistive (fig. 12*a*), the positions of the maxima and minima do not depend very much upon $|r|$, and the value, represented by $|r'|$, deduced from the ripple ratio $a_r/a_0$, does not differ appreciably from the correct value $|r|$ (fig. 12*a* gives 0.30 and 0.77 instead of 0.33 and 0.80, respectively).

However, as soon as $Z_g$ acquires a reactive component, the maxima and (to a smaller extent) the minima are displaced. This is owing to the fact that in the case of, say, a capacitative $Z_g$, the voltage $|V(l)|$ does not reach a maximum when the input impedance of the cable is real, but does do so in the case of an inductive input impedance, where a build-up takes place.

To calculate $\varphi$ from equation (12), then, it is most convenient to deduce the difference in frequency from the interval between minima, since they are more stable than the maxima. The effect of the displacement can be further reduced by adding to the complex impedance $Z_g$ a reactance such that the two combined assume a real value (the capacitative impedance of the detector in parallel with $Z_g$ is always a part of $Z_g$).

*Effect of the detector characteristic*

*Fig. 13* shows the relationship between the amplitude ($V$) of the high-frequency voltage, and the detected D.C. voltage ($U$), in the case of a germanium diode [7]) (this being more suitable than a vacuum diode by virtue of the fact that its capacitance is lower). Usually, owing to the non-linearity of this characteristic, the variation of voltage $U$ does not entirely correspond to the changes in $V$.

With the method considered here, the oscilloscope is employed to observe the fluctuation of $U$, instead of $V(l)$. Here, then, the ratio of the ripple amplitudes of $U$ as derived from the oscillogram will differ from that of the ripple amplitudes of $|V(l)|$, which is required for the determination of $|r|$ from (10).

The error arising from this difference in ratio has been calculated in one or two cases. With linear detection, and with

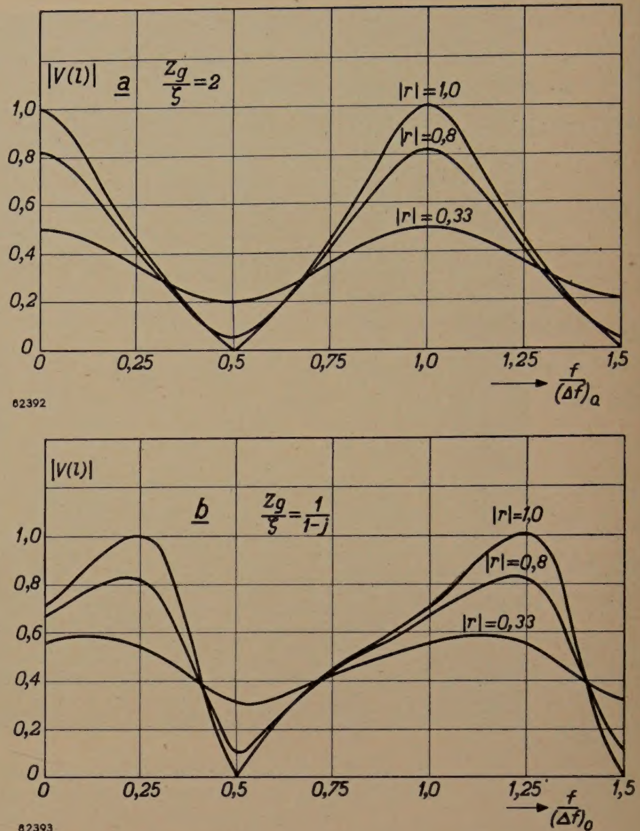


Fig. 12. $|V(l)|$ plotted against $f$ when $Z_g$ does not equal $\zeta$. Curves ($a$) refer to $Z_g = 2\zeta$ (real), and curves ($b$) to $Z_g = \zeta/(1-\mathrm{j})$; in both cases $|r|$ has the values 0.33, 0.80 and 1.0. In ($a$) the position of the maxima and minima is virtually independent of $|r|$. In ($b$), the position of the maxima is largely determined by $|r|$; for the minima this is not so. It is assumed that $E$ is unity.

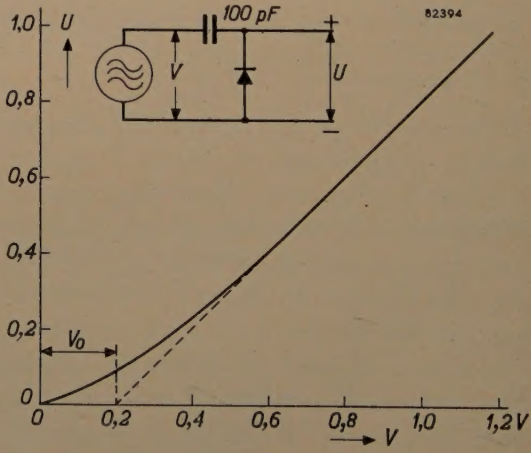


Fig. 13. Relationship between the detected D.C. voltage ($U$) and the amplitude ($V$) of the high-frequency A.C. voltage, for a given germanium diode. The value ($V_0$) at which the straight part of the characteristic (produced) would cut the $V$ axis is about 0.2 V.

---

[7]) See Philips tech. Rev. **16**, 225-232, 1954/55 (No. 8).

square-law detection there is no error. These cases correspond roughly to $E \gg V_0$ and $E \ll V_0$, respectively (for $V_0$ see fig 13). However, the output voltages of most wobbulators (0.1-0.5 V) are of the same order of magnitude as $V_0$; hence we find in the characteristic of fig. 13 a deviation corresponding to line $b$ in *fig. 14.*

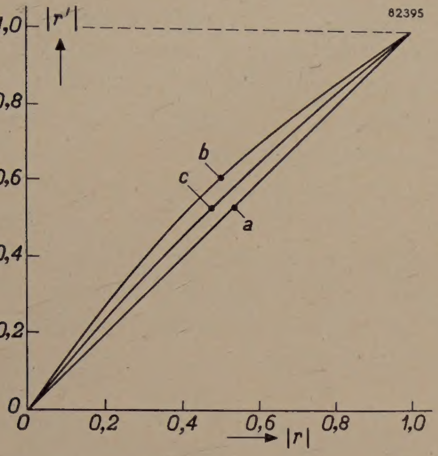


Fig. 14. Given the detection characteristic shown in fig. 13, and with $E$ of the same order of magnitude as $V_0$, the measured value of the reflection coefficient will be too high ($|r|'$); note the deviation of curve $b$ with respect to the straight line at an angle of 45°. Attenuation reduces this error, as will be seen from the curve ($c$) referring to $e^{-2al} = 0.8$.

Accordingly, it is expedient to increase the attenuation of the line, since with total reflection, the minima of $|V(l)|$ then remain above zero (fig. 11$b$) and therefore cover a relatively smaller portion of the bend in the detection characteristic. If $e^{-2al}$ is say, 0.8, the deviations will be halved (curve $c$ in fig. 14).

## Applications

### *Matching a tuned circuit to a cable*

The impedance $Z$ of a parallel circuit ($L$, $C$, and $R$ in parallel) may be written as:

$$Z = \frac{R}{1 + \mathrm{j}\beta_0 Q},$$

where $\beta_0 = \frac{\omega}{\omega_0} - \frac{\omega_0}{\omega}$ (with $\omega_0^2 = 1/LC$) is the detuning (here indicated by $\beta_0$ to distinguish it from the phase constant $\beta$), and $Q$ the quality factor, i.e. $R\sqrt{C/L}$. From (5), when the circuit is connected to the transmission line, the reflection coefficient is:

$$r = \frac{\dfrac{R}{1+\mathrm{j}\beta_0 Q} - \zeta}{\dfrac{R}{1+\mathrm{j}\beta_0 Q} + \zeta}.$$

Hence the modulus of $r$ is:

$$|r| = \sqrt{\frac{(R-\zeta)^2 + \beta_0^2 Q^2 \zeta^2}{(R+\zeta)^2 + \beta_0^2 Q^2 \zeta^2}}. \quad . \quad . \quad (13)$$

If the frequency be so increased or reduced that $\omega$ passes resonance ($\omega_0$), $\beta_0^2$ will drop to zero, and then rise again: from (13), $|r|$ has a minimum $(R-\zeta)/(R+\zeta)$ at resonance. When circuit and cable are matched ($R = \zeta$), this minimum is zero.

As we have already seen, the voltage $V(l)$ at the input of the transmission line is equivalent to the vector sum of the stationary vector $\frac{1}{2}E$ and the rotating vector $|r| \cdot \frac{1}{2}E$. As we have seen above, $|r|$ varies during the frequency modulation cycle, that is, decreases to begin with, but increases when the frequency goes beyond resonance. In response to this variation, the end of vector $V(l)$ first describes an inward spiral, and then an outward spiral. Accordingly, the oscillogram shows, from left to right, an initial decrease in amplitude followed by an increase. When the minimum is zero, line and circuit are matched; this enables us to see at a glance whether they are matched or not.

In the case of a mismatch, matching can be effected by connecting the line to a tapping in the circuit (*fig. 15a*). It is then necessary to substitute $R/\tau^2$ for $R$ in formula (13), $\tau$ being the transformation ratio. Fig. 15 depicts $|r|$ plotted against $\beta_0 Q$ for various values of $\tau$; matching occurs where $\tau = \tau_{\mathrm{opt}} = \sqrt{R/\zeta}$.

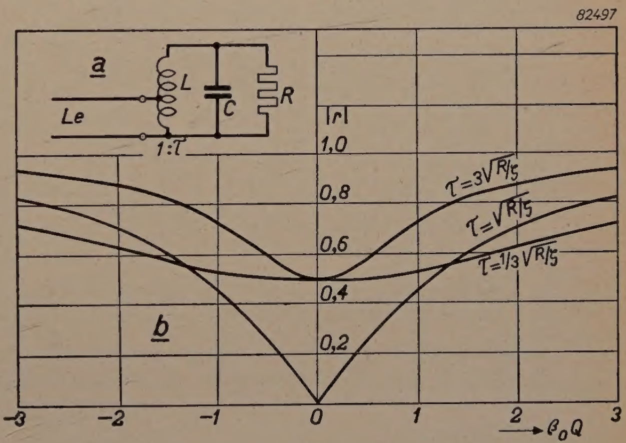


Fig. 15. $a$) The impedance of a resonant circuit ($L$, $C$, $R$) can be matched to that of the transmission line ($Le$) by employing a suitable transformation ratio ($\tau$).
$b$) Relationship between $|r|$ and the detuning $\beta_0$ (the quantity actually plotted is $\beta_0 Q$), for various values of the transformation ratio ($\tau$). The optimum ratio ($\tau_{\mathrm{opt}}$) is $\sqrt{R/\zeta}$.

*Fig. 16* shows the results of experiments with three different tappings on the coil of a circuit tuned to about 560 Mc/s. Fig. 16$b$, where the minimum amplitude is about zero, is the closest approximation to ideal matching. The upper three envelopes in fig. 16 correspond to the curves shown in fig. 15$b$ [8]).

---

[8]) It should be noted that although this variation is related to the quality factor $Q$, the value of $Q$ cannot be deduced from it as readily as from a resonance curve.

Both components of the impedance are derived from the measured values of the reflection coefficient. It can be shown (although we must omit the proof) that in the Smith chart the measuring points corresponding to the different frequencies should form a circle. The extent to which this condition is satisfied will be seen from *fig. 17*.

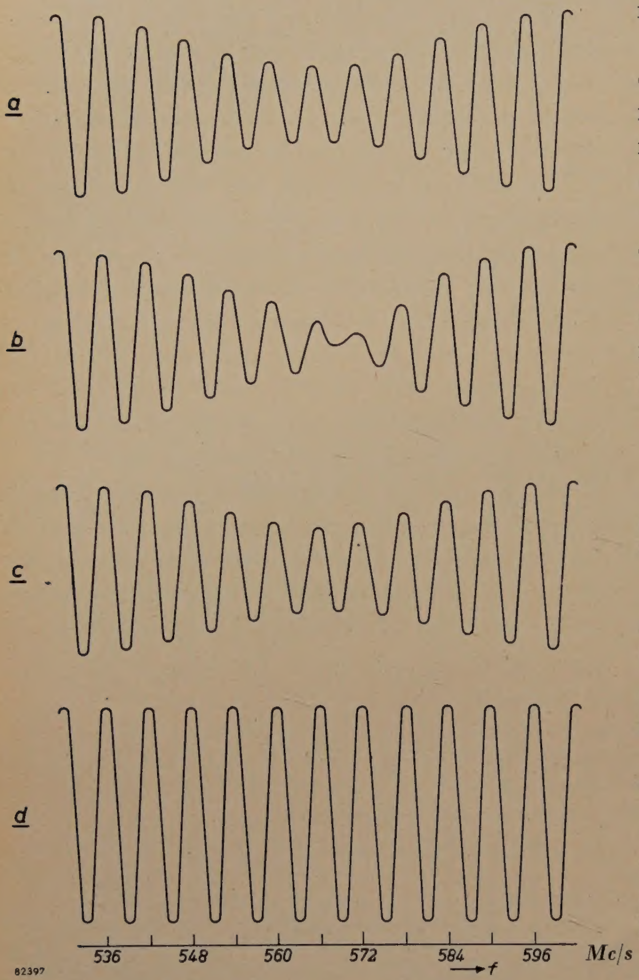


Fig. 16. Experimental results obtained in a case similar to that illustrated in fig. 15*a*. Here, the circuit is tuned to about 560 Mc/s.
*a*) $\tau$ too high, *b*) matching virtually correct, *c*) $\tau$ too low. In case (*d*), the end of the transmission line is short-circuited.

### *Measuring cable characteristics*

The method considered here is also very effective as a means of measuring cable characteristics.

Firstly, let us consider the characteristic impedance $\zeta$. This can be measured by terminating the cable with such a resistance that the oscillogram has a constant amplitude. Although in this case the generator impedance $Z_g$ cannot be matched to $\zeta$ (this being unknown), it does not detract from the accuracy of the measurement, since, provided that $R = \zeta$, the oscillogram will show a constant amplitude even if $Z_g \neq \zeta$. At the frequencies involved, continuously variable resistors of the required order of magnitude (some tens or hundreds of ohms) are not sufficiently free of self-inductance and capacitance; therefore, a number of fixed resistors of known value are employed instead.

The reflection coefficient $r$, is measured with the cable connected to each resistor in turn, and the resistance corresponding to $|r| = 0$ is then determined by interpolation; this resistance equals the characteristic impedance.

Another cable-constant, $k = v/c = 1/\sqrt{\varepsilon_{\mathrm{rel}}}$ (where $c$ is the velocity of light and $\varepsilon_{\mathrm{rel}}$ the relative dielectric constant), may be defined as the factor expressing the wavelength in in the cable as a fraction of the corresponding wavelength in air. From (11) we have:

$$k = \frac{2l(\Delta f)_0}{c}.$$

$(\Delta f)_0$ can be determined accurately by means of the "marker" already described.

By this method it was found that for coaxial and shielded twin cables, with solid polythene insulation $k = 0.65$. For unshielded ribbon type twin lead (likewise insulated with polythene), of characteristic impedance $\zeta = 300$ ohms, a usual value is $k = 0.79$.

Lastly, let us consider the attenuation in a cable. The amplitude of the oscillogram is proportional to $E(1 - \mathrm{e}^{-2\alpha l})$ when the end of the cable is short-circuited, and to $E$ when the cable is omitted altogether. From these amplitudes, $\alpha$ can be calculated.

### *Determining the effect of terminal insulators*

TV and F.M. receivers are often connected to a roof-aerial by means of an unshielded polythene twin lead, which must be secured at various points by means of insulators. The insulators impair the uniformity of the cable and so cause reflections, which may affect the reception of the signal. Even a relatively small number of such irregularities may give rise to a most confused situation; hence reflections should be minimized as far as possible.

The measurements that will now be described illustrate a simple case. A long ribbon type twin lead is terminated with a resistor matching it as far as possible and thus ensuring a low reflection coefficient. Now, curve *a* in *fig. 18* shows measured values of $|r|$ plotted against the frequency ($f$) up to about 1000 Mc/s. An insulator is attached to the cable, close to the terminating resistor. Measured values of $|r|$ are again plotted as a function of $f$ (the total reflection required for this purpose being procured by short-circuiting the cable just beyond the insulator). The result shows a considerable increase in the reflection coefficient (curve *b* in fig. 18). In the case considered it was shown that the effect of the insulator was equivalent to a capacitance of 0.13 pF between the cable cores at the point of attachement.

Another insulator is now fixed to the cable, about 7.5 cm from the first. This roughly doubles the reflection coefficient at the lower frequencies (about 100 to 300 Mc/s, curve *c* in fig. 18), but reduces it considerably at frequencies in the region of 800 Mc/s. This is understandable in view of the fact that at 800 Mc/s the wavelength in the cable is $kc/f = 0.8 \times 3 \times 10^8/(800 \times 10^6) = 0.30$ m; hence the distance travelled by the wave from one insulator to the other and back again is precisely half a wavelength and the waves will therefore cancel one another.

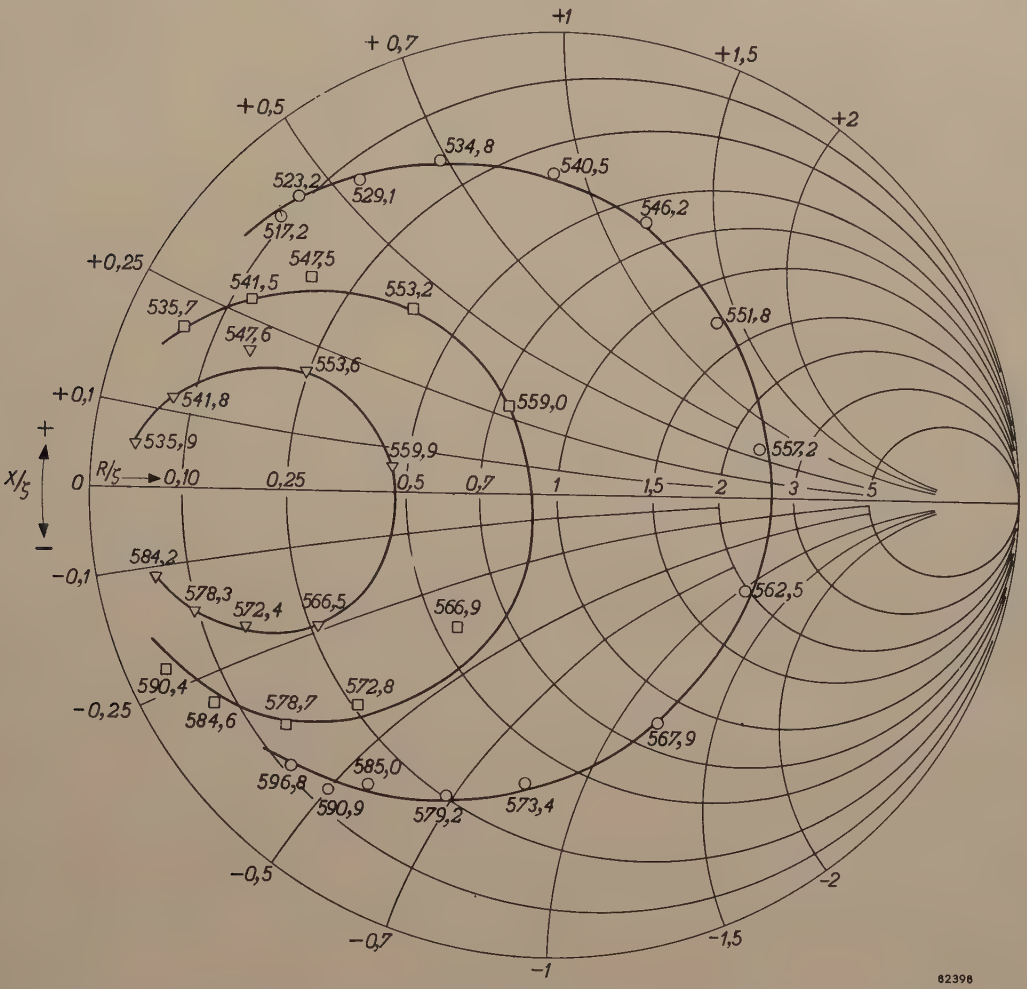


Fig. 17. Impedances as measured in the cases illustrated in fig. 16*a*, *b* and *c*, plotted in Smith's diagram. Theoretically the curves referring to a constant transformation ratio ($\tau$) should be circles. The appropriate frequency (in Mc/s) is indicated at each point.

**Variants of the method described**

The wobbulator and oscilloscope method is very quick and convenient in that it provides an immed-

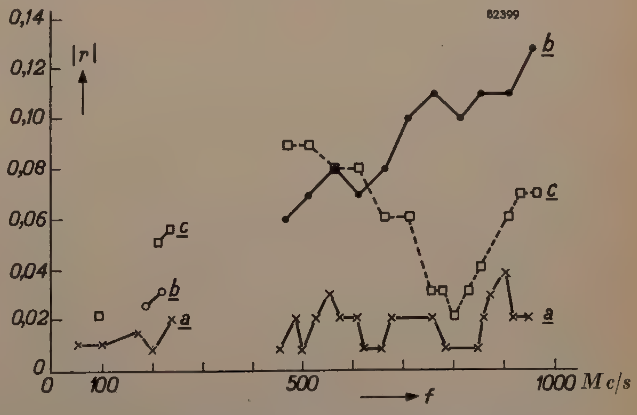


Fig. 18. Measured values of the modulus $|r|$ of the reflection coefficient in a unshielded polythere twin lead with $\zeta = 300$ ohms, plotted against the frequency *f*. The cable is matched fairly well by a resistance. *a*) Without deliberately introduced irregularities, *b*) with one insulator close to the terminating resistor, *c*) with another insulator about 7.5 cm from the first.

iate indication of the effect of any modification to the object under test. In fact, this method is most suitable for experimental measurements covering a wide range of frequencies.

On the other hand, the results plotted from an oscillogram are never as accurate as those obtained with direct-reading instruments, Where accurate numerical results are required, the following procedure may be adopted. A simple variable-frequency oscillator is substituted for the wobbulator and the D.C. current from the detector is measured with a microammeter. The frequencies associated with the maximum and minimum current values are determined, and these values themselves are read from the meter. The measurement is then carried out once more with the line either open, or short-circuited. Lastly, $|r|$ and $\varphi$ are calculated with the aid of formulae (10) and (12), as before.

Better still is to employ an oscillator connected to a calibrated attenuator. The latter is then so adjusted as to maintain the D.C. detector current constant (with the object of avoiding any variation in the detector load, so that the detector characteristic does not affect the measurement). The ripple amplitudes are then computed from the attenuator readings.

Another refinement is to modulate the oscillator in *amplitude* at a low frequency. The detector then produces a low-frequency voltage, which can be measured with a very sensitive electronic voltmeter. This method enables a low input voltage to be employed; any valves in the object under test will not then be overloaded.

**Summary.** An impedance can be measured by employing it as the terminating impedance of a transmission line. The transmission line in the form of a cable about 60 m long, is fed from a variable-frequency oscillator. A detector (germanium diode) with an indicating instrument is connected to the line input. The oscillator frequencies corresponding to maximum and minimum deflections in the indicator are then determined, once with the line terminated with the unknown impedance, and once with the end of the line either short-circuited or open (total reflection). The argument and modulus of the (complex) reflection coefficient are then calculated from these frequencies and the associated readings, and used to plot the unknown impedance in the Smith chart.

One of the practical forms of this method uses a frequency-modulated oscillator — that is, a wobbulator — and an oscilloscope, whose horizontal deflection is synchronized with the frequency modulation. The oscillograms then provide data concerning the reflection coefficient, from which the impedance can be deduced; moreover, they show at once whether the impedance and the line are matched or not.

One or two applications are discussed, e.g. the matching of a tuned circuit to a cable, the measurement of various cable characteristics, and the determination of the effect of irregularities arising from insulators on an unscreened twin lead.

The wobbulator and oscillograph method is most suitable for quick experimental measurements covering a wide range of frequencies. For more accurate numerical results it is better to substitute a variable oscillator (preferably amplitude-modulated), with a calibrated attenuator, for the wobbulator, and a D.C. microammeter (or electronic millivoltmeter) for the oscilloscope.

# ENTROPY IN SCIENCE AND TECHNOLOGY

## III. EXAMPLES AND APPLICATIONS (continued)

by J. D. FAST.

536.75

*As in the previous article (II), the author demonstrates with examples the important part played by entropy in widely differing fields of science and technology. This article deals with the theory of the specific heat of solids, the thermodynamic equilibrium of certain lattice defects in solids and their relationship with diffusion phenomena, and the equilibria of mixtures, both in a solid form (e.g. alloys) and in the form of liquid solutions, including the extreme case of solutions of polymers. The statistical aspects of rubber elasticity are also discussed, a phenomenon that can be completely described as an entropy-effect.*

The first article of this series [1]), referred to below as I, gave some general observations on the concept of entropy; the second [2]), referred to as II, presented some applications of this concept in chemistry, physics and technology. In the present, third article, some further examples are discussed. A fourth article, to be published later, will be of a somewhat different nature, and will discuss the application of the concept of entropy to information theory.

### The specific heat of an Einstein solid

In I (page 262) we discussed a simplified model of a solid, in which the identical atoms behave as linear harmonic oscillators, executing their vibrations around fixed centres practically independently of one another, and capable of absorbing equal quanta of a value $h\nu$. Suppose that this solid, consisting of $N$ atoms, absorbs a number $q$ of energy quanta $h\nu$ (starting from absolute zero). This energy can be distributed among the oscillators in a great number of ways (it is assumed that $N \gg 1$ and $q \gg 1$). Each distribution represents one microstate and the total number of micro-states $m$ is given, according to (I, 9) by:

$$m = \frac{(q+N-1)!}{q!\,(N-1)!}.$$

For the sake of convenience it was not taken into account when deriving this formula that an atom in an Einstein solid has to be assigned *three* degrees of vibrational freedom. An actual crystal of $N$ atoms in this model can be regarded as a system of $3N$ linear oscillators. Since real crystals as a rule contain more than $10^{19}$ or $10^{20}$ atoms, unity is negligible compared to $N$, and the formula becomes:

$$m = \frac{(q+3N)!}{q!\,(3N)!}\,. \quad . \quad . \quad . \quad . \quad (\mathrm{III}, 1)$$

The energy of the crystal, after $q$ vibrational quanta have been absorbed, has risen to an amount

$$U = q\,h\nu \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\mathrm{III}, 2)$$

above the zero-point energy, whilst the entropy, according to (I, 12) and (III, 1) and using Stirling's formula in the approximation (I, 3) is given by

$$S = k\,\{(q+3N)\ln(q+3N) - q\ln q - 3N\ln 3N\}. \quad (\mathrm{III}, 3)$$

The Helmholtz free energy, $F = U - TS$, may therefore be written as

$$F(q) = qh\nu - kT\,\{(q+3N)\ln(q+3N) - \\ - q\ln q - 3N\ln 3N\}. \quad (\mathrm{III}, 4)$$

Until now the number of absorbed quanta has been assumed to be known. In fact it is usually not primarily this number of quanta which is known, but the temperature to which the crystal is heated by bringing it into thermal contact with surroundings at constant temperature. What we wish to know, therefore, is the number of quanta $q$ present at a given temperature $T$ in the Einstein solid. The knowledge of $q$ as a function of $T$ will enable us to calculate the specific heat at any temperature.

If the solid could entirely submit to its "striving" towards a minimum value of of the energy, it would not absorb any quanta at all. Conversely, if it could entirely submit to its "striving" towards a maximum value of the entropy, the number of quanta absorbed would continue to increase. The compromise (the state of equilibrium) lies, according to I, at the point where the Helmholtz free energy is a minimum, i.e. where

---

[1]) J. D. Fast, Entropy in science and technology, I, The concept of entropy, Philips tech. Rev. 15, 258-269, 1954/55.

[2]) J. D. Fast, Entropy in science and technology, II, Examples and applications, Philips techn. Rev. 16, 298-308, 1954/55.

$$\frac{\mathrm{d}F(q)}{\mathrm{d}q}=0.$$

Differentiating (III, 4) and equating to zero,

$$h\nu - kT\ln\frac{q+3N}{q}=0$$

or, after re-arranging,

$$q=\frac{3N}{\mathrm{e}^{h\nu/kT}-1}. \quad . \ . \ . \ . \ \text{(III, 5)}$$

The specific heat at constant volume is given by:

$$c_v=\frac{\mathrm{d}Q}{\mathrm{d}T}=\frac{\mathrm{d}U}{\mathrm{d}T}=\frac{\mathrm{d}(qh\nu)}{\mathrm{d}T}.$$

Substituting for $q$ from (III, 5) gives the relationship sought:

$$c_v=3\,kN\left(\frac{h\nu}{kT}\right)^2\frac{\mathrm{e}^{h\nu/kT}}{(\mathrm{e}^{h\nu/kT}-1)^2}. \qquad \text{(III, 6)}$$

This formula was deduced by Einstein in 1907. For the part of the internal energy that varies with temperature we can write, according to (III, 2) and (III, 5):

$$U=\frac{3N\,h\nu}{\mathrm{e}^{h\nu/kT}-1}. \quad . \ . \ . \ . \ \text{(III, 7)}$$

If $kT$ is very much greater than $h\nu$, then

$$\mathrm{e}^{h\nu/kT}\approx 1+h\nu/kT$$

and therefore,

$$U\approx 3\,NkT,$$

$$c_v\approx 3\,Nk.$$

At relatively high temperatures, therefore, the specific heat of an Einstein solid reaches a constant value which is not only independent of the temperature but also of $\nu$. If this idealized solid represented the behaviour of actual (elementary) solids, then the latter would all have the same specific heat per gram-atom at not too low temperatures. This would amount to $c_v = 3\ N_0k = 3R$, in which $N_0$ represents Avogradro's number and $R$ the gas constant. According to the experimentally established law of Dulong and Petit, many solid elements do in fact have an approximately equal specific heat of about 6 cal per degree per gram-atom ($R$ = approx. 2 cal per degree) at not too low temperatures It is also in agreement with observation that (III, 6) requires that $c_v$, at decreasing temperature, finally approaches zero; cf. *fig. 1.* which represents the relationship between $c_v/R$ and $kT/h\nu$ according to (III, 6).

To a first approximation, Einstein's formula thus provides a good expression for the thermal behaviour of the solids considered here. The form of the experimental $c_v(T)$ curves does not agree in detail with the formula, however. The largest deviations occur at very low temperatures. This is due to the fact that a real crystal cannot, in fact, be considered as an asssembly of nearly independent oscillators of equal frequency as assumed in the Einstein derivation. A crystal has a great number of vibration modes at widely divergent frequencies; they may not be regarded as the vibrations of individual atoms, for they are intrinsic to the lattice as a whole: they can be pictured as a pattern of standing waves in the crystal. A theory based on this model and in better agreement with experiment has been formulated by Debye. We shall not enter into this theory in view of the fact that, considered purely thermodynamically, it covers no fresh ground. One of its consequences, which is in good agreement with experiment, may be mentioned: at low temperatures the Stefan-Boltzmann law (cf. II) is applicable to the vibrations in a solid. According to this law, the energy is proportional to $T^4$, so that the specific heat is proportional to $T^3$. This consequence is due to the fact that both the radiation inside a hollow body and the vibrations in a solid can be considered as a broad spectrum of modes of vibration. It is true that the hollow space has an infinite number of modes, whereas the number of modes in the solid is restricted by the number of atoms $N$ (it amounts to $3N$). At low temperatures, however, this difference is irrelevant since then, according to quantum mechanics, the high-frequency modes play no part.

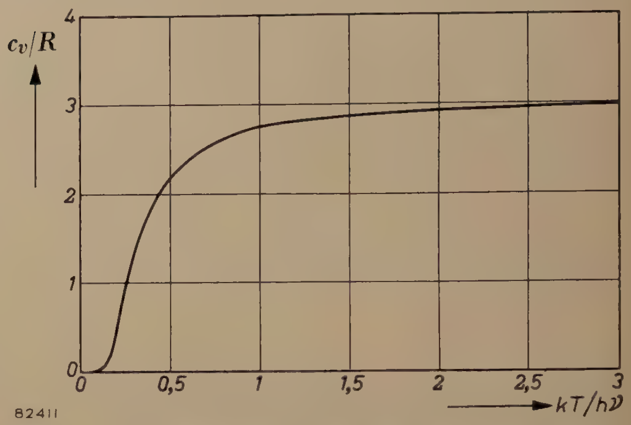


Fig. 1. $c_v/R$ as a function of $kT/h\nu$ for an Einstein solid (formula (III, 6)).

We have already seen that to consider a solid as an assembly of independent oscillators of equal frequency is rather an inexact model. A diatomic or polyatomic gas, on the other hand, conforms very satisfactorily to this model. In a diatomic gas each molecule can vibrate in such a manner that the atoms oscillate along their line of centres. (Translational and rotational degrees of freedom also contribute to the specific heat of a gas, but will not be discussed here.) Exchanges of energy are brought about by collisions between the gas molecules. Apart from this the vibrations are mutually independent. Formula (III, 6), with the omission of the factor 3, is therefore applicable with considerably greater accuracy to the vibrations of a diatomic gas than to those of a solid. The specific-heat values calculated on this basis provide the necessary corrections to give validity to the calculation of chemical equilibria (see the beginning of II) even at very high temperatures.

### Vacancies and diffusion in solids

Atomic diffusion plays an important part in several processes occurring in solids. Diffusion in metals and alloys in particular has been the subject of extensive research during the last few decades. This has revealed that the presence of defects in the periodic crystal lattice, particularly vacancies or interstitial atoms, is essential for the occurrence of diffusion.

As we have seen in the foregoing discussion of an Einstein solid, the vibrations of the atoms around their positions of equilibrium become more and more violent as the temperature rises. Even before the melting point is reached, a fraction of the atoms will possess sufficient energy to leave their lattice positions completely. These atoms form new lattice planes on the outside of the crystal, since the spaces between the other atoms, the interstices, are too small to accommodate them. Starting, then, with a perfect crystal (no defects), any vacancies which occur in the crystal as the temperature rises must originate in the boundary layer. One can imagine that a few atoms in the boundary layer leave their original positions and occupy new positions on the surface (*fig. 2*). Atoms from deeper layers can subsequently jump into the newly created vacancies, and so on. We may equally well say that the vacancies arise at the surface and subsequently diffuse to the interior. The fact that at high temperatures lattice defects are bound to arise even in the state of equilibrium, is due to the fact that their occurrence represents an increase in the entropy. Disregarding the extremely small macroscopic volume changes occurring with the formation of the vacancies, it can be said that at a certain temperature $T$ vacancies will continue to be formed until ultimately the Helmholtz free energy $F = U - TS$ has reached its minimum value. Although the internal energy $U$ will rise by an amount $\Delta U$ due to the introduction of vacancies, it is nevertheless possible that under certain conditions the value of the free energy will drop, viz. if $T\Delta S > \Delta U$.

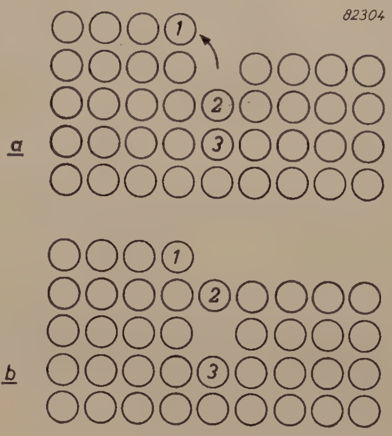


Fig. 2. The vacancies in a solid are assumed to form at the outer surface (*a*) and subsequently diffuse to the interior (*b*), this being equivalent to the outward diffusion of atoms (*1*, *2*, *3*).

If the energy $\varepsilon$ required for forming a vacancy is known, the percentage of vacancies can be evaluated for any temperature. This is done as follows.

If a perfect crystal of $N$ atoms changes into an imperfect state with $n$ vacancies, then the corresponding increase in entropy, $\Delta S = k \ln m$, is given by the number of different ways $m$ in which $N$ identical atoms can be accommodated in a lattice with $(N + n)$ available positions. This number is

$$m = \frac{(N+n)!}{N!\,n!}. \quad . \; . \; . \; . \quad \text{(III, 8)}$$

Using Stirling's formula in the approximation (I, 3), we find:

$$\Delta S = k\{(N+n)\ln(N+n) - N\ln N - n\ln n\}.$$

As long as the concentration of the vacancies is so small that their interaction is negligible, the increase in the energy is determined by $\Delta U = n\varepsilon$ and thus the change of the free energy by

$$\Delta F = \Delta U - T\Delta S =$$
$$= n\varepsilon - kT\{(N+n)\ln(N+n) - N\ln N - n\ln n\}. \quad \text{(III, 9)}$$

In the equilibrium state, the free energy is a minimum; thus $\partial F/\partial n = 0$, or, from (III, 9):

$$\varepsilon - kT\{\ln(N+n) - \ln n\} = 0,$$

i.e. $$\frac{n}{N+n} = \mathrm{e}^{-\varepsilon/kT}.$$

Since $n$ is negligible with respect to $N$, this becomes:

$$\frac{n}{N} = \mathrm{e}^{-\varepsilon/kT} = \mathrm{e}^{-E/RT}, \quad . \;.\; . \quad \text{(III, 10)}$$

where $E$ is the energy required to make $6 \times 10^{23}$ (Avogadro's number) vacancies. According to the calculations of Huntington and Seitz [3]) $\varepsilon$ has the value of approximately 1 eV for copper, i. e. $E \approx$ 23 000 cal per "gram-atom of vacancies". For 1000 °K, therefore, we arive at:

$$\frac{n}{N} = \mathrm{e}^{-23\,000/2000} = 10^{-5}.$$

In this case one in each 100 000 lattice positions would be unoccupied. This corresponds to an average distance between adjacent vacancies of the order of thirty interatomic distances.

The additional contribution to the internal energy owing to the formation of vacancies is accompanied by an additional contribution to the specific heat, given by:

$$\frac{\mathrm{d}\Delta U}{\mathrm{d}T} = \frac{\mathrm{d}n\varepsilon}{\mathrm{d}T} = \frac{\mathrm{d}(N\varepsilon\mathrm{e}^{-\varepsilon/kT})}{\mathrm{d}T} = R\left(\frac{\varepsilon}{kT}\right)^2 \mathrm{e}^{-\varepsilon/kT} \quad \text{per gram-atom.}$$

The above is based on the assumption that the increase in the entropy during the transition from the ideal to the disturbed arrangement is exclusively based on the number of possible distributions $m$, given by formula (III, 8). In reality the vibration frequencies of the atoms near the vacancies also change; this contributes to the entropy.

This may be explained as follows. An atom adjacent to a vacancy is bound more weakly than an atom completely surrounded by similar atoms. Consequently it will have a lower vibration frequency, at least in the direction of the vacancy. Since, from (III, 3) and (III, 5).

$$S \approx 3\,Nk \ln \frac{kT}{h\nu} \quad \text{for } kT \gg h\nu, \quad . \;. \quad \text{(III, 11)}$$

is applicable to an Einstein solid, the introduction of vacancies causes not only the entropy increase as evaluated in (III, 8), but also an increase in the vibrational entropy. The concentration of the vacancies can, therefore, be considerably greater than the value calculated above.

At high temperatures the vacancies move at random through the lattice. A displacement of a vacancy over one interatomic distance is brought about if an adjacent atom jumps into the vacant site. For such a jump an amount of energy, at least equivalent to the activation energy $q$ is required. This energy will, as a rule, be large compared to the mean thermal energy of the atoms. The fraction of the total number of atoms which possesses at least the activation energy $q$ is given by

$$f = \mathrm{e}^{-q/kT} = \mathrm{e}^{-Q/RT},$$

where $Q = N_0 q$ is the "activation energy per gram-atom". The diffusion constant $D$ of the *vacancies* will be approximately proportional to this factor:

$$D = D_0\,\mathrm{e}^{-Q/RT}, \quad . \;.\; .\; . \quad \text{(III, 12)}$$

where $D_0$ is a constant which is independent or only slightly dependent on the temperature.

Primarily, however, we are often not so much interested in the diffusion of the vacancies as in the diffusion of the atoms of the pure metal, i.e. the self-diffusion, or the diffusion of foreign atoms occupying the lattice sites ( substitutional atoms). The self-diffusion can be studied with the aid of certain radioactive isotopes of the atoms of the pure metal. The diffusion constant for this type of diffusion should be smaller than that for the vacancy diffusion, because the atoms are capable of jumping only at that moment when there happens to be a neighbouring vacancy. The number of times per second that this condition occurs is proportional to the diffusion constant of the vacancies and to their relative number, and hence to (III, 12) and (III, 10).

According to this reasoning, the diffusion coefficient will be given by an expression of the form:

$$D = D_0{}'\,\mathrm{e}^{-Q/RT}\,\mathrm{e}^{-E/RT}$$

or

$$D = D_0{}'\,\mathrm{e}^{-W/RT}, \quad . \;.\; .\; . \quad \text{(III, 13)}$$

where

$$W = Q + E \quad . \;.\; .\; .\; .\; . \quad \text{(III, 14)}$$

and $D_0{}'$ is a constant which is independent or only slightly dependent upon the temperature.

The foregoing also applies to substitutional atoms, provided that the properties and size of these atoms differ so little from those of the solvent that the vacancies have no preference for either type of atom, and that the necessary activation energy is the same for both.

## Alloys

### *Entropy of mixing and energy of mixing*

An enormous number of different alloys with a great diversity of properties can be made, owing to the almost limitless combinations which are possible: many metallic elements can be combined in several ways into binary, ternary, etc. alloys, the relative quantities in each combination can be varied, and each alloy can be subjected to a great variety of heat and mechanical treatments.

The central problem in the metallurgy of alloys is that of the solubility of the different metals, more particularly the solubility in the solid state. If we confine ourselves to the alloys of two metals (binary alloys), we find that the majority of combinations, no matter in what proportion they are mixed, form a homogeneous mixture in the liquid state, whereas in the solid state they have only a limited miscibility. Miscibility in all proportions in only possible in the solid state if both metals have the same crystal structure and if their atoms show only slight differences in size and electron configuration.

---

[3]) H. B. Huntington and F. Seitz, Phys. Rev. **61**, 315-325, 1942.

Homogeneous mixing is in nearly all cases accompanied by an increase of the entropy, i.e. it is nearly always promoted by the entropy effect. The separation of the mixture into two phases can only occur if also the energy increases during the homogeneous mixing. If the energy does not change or even decreases in the course of mixing, then only one homogeneous phase can exist in the state of equilibrium. The entropy increase occurring when two metals $A$ and $B$ are homogeneously mixed, can again be directly established by a statistical reasoning, provided that all configurations are equally probable. If we consider a total number of $N$ metal atoms, of which $Nx$ are of type $B$ and hence $N(1-x)$ are of type $A$, then the number of possible ways in which these can be distributed among the available lattice points is given by:

$$m = \frac{N!}{(Nx)!\,\{N(1-x)\}!} \quad . \; . \; . \quad \text{(III, 15)}$$

The increase in entropy that accompanies the mixing i.e. the entropy of mixing $\Delta S = k \ln m$, thus becomes:

$$\Delta S = Nk \{-x \ln x - (1-x) \ln (1-x)\}. \quad \text{(III, 16)}$$

Strictly speaking, this expression only gives the entropy of mixing at 0 °K. If it is assumed that the vibrational entropy (cf. III, 11) changes little or not all in the course of mixing, (III, 16) is also applicable at other temperatures.

To find a mathematical expression for the energy of mixing (also called the heat of solution) we start from the over-simplified assumption that the internal energy at zero temperature can be written as the sum of the binding energies of nearest neighbours only. As a consequence of this only three interaction energies $\varepsilon_{AA}$, $\varepsilon_{BB}$ and $\varepsilon_{AB}$ between neighbouring pairs $A$-$A$, $B$-$B$ and $A$-$B$ will occur in the terms of this sum [4]). In order to do the summation we have to evaluate the numbers of the three types of neighbour configurations.

The number of nearest neighbours to any one atom in alloy is denoted by $z$ ($z = 8$ for the body-centred cubic structure and $z = 12$ for the close-packed cubic and hexagonal structures). If the atoms of the type $A$ and $B$ are distributed at random among the lattice points, then the chance of an atom $A$ occupying any given lattice point is $(1-x)$ whilst that of an atom $B$ is $x$. The probability of finding the combination $A$-$A$ in two adjacent positions is further given by $(1-x)^2$, that of finding $B$-$B$ by $x^2$ and that of finding either $A$-$B$ or $B$-$A$ by $2x(1-x)$. In total there are $Nz/2$ bonds between the $N$ atoms of the alloy; in this particular case they are distributed as follows:

$$\frac{Nz}{2}(1-x)^2 \quad A\text{-}A \text{ bonds,}$$

$$\frac{Nz}{2}x^2 \qquad B\text{-}B \text{ bonds,}$$

$$Nz\,x\,(1-x) \quad A\text{-}B \text{ bonds.}$$

The energy of the alloy at 0 °K is therefore given by:

$$U_{AB} = \frac{Nz\varepsilon_{AA}}{2}(1-x)^2 + \frac{Nz\varepsilon_{BB}}{2}x^2 + Nz\varepsilon_{AB}x\,(1-x). \quad \text{(III, 17)}$$

For pure $A$ ($x = 0$) the formula gives:

$$U_{AA} = \frac{Nz\,\varepsilon_{BB}}{2};$$

for pure $B$ ($x = 1$):

$$U_{BB} = \frac{Nz\,\varepsilon_{BB}}{2}.$$

For a *heterogeneous mixture* of the pure metals, the internal energy is given by:

$$\frac{Nz\varepsilon_{AA}}{2}(1-x) + \frac{Nz\,\varepsilon_{BB}}{2}x. \quad . \; . \quad \text{(III, 18)}$$

The heat of solution at 0 °K is given by the difference between (III, 17) and (III, 18):

$$\Delta U = x(1-x)\,Nz\left\{\varepsilon_{AB} - \frac{\varepsilon_{AA}+\varepsilon_{BB}}{2}\right\}. \quad \text{(III, 19)}$$

As a rule the vibrational energy will not have much influence during the mixing process, so that (III, 19) is also valid at higher temperatures.

Since $x(1-x)Nz$ is always positive, the sign of $\Delta U$ is determined by that of $\varepsilon_{AB} - \frac{1}{2}(\varepsilon_{AA} + \varepsilon_{BB})$.

### *Solution, ordering, segregation and precipitation*

If in the foregoing $\varepsilon_{AB}$ were to equal the mean of $\varepsilon_{AA}$ and $\varepsilon_{BB}$ (which will probably never be exactly true), then $\Delta U = 0$. As regards energy, there is no preference whatsoever for a specific bond, and the entropy effect causes a homogeneous mixing in the state of equilibrium.

---

[4]) We would ultimately find the same formula (III, 19) — still to be derived — if we also took into account, the interaction between *more remote* pairs. The energies $\varepsilon_{AA}$, $\varepsilon_{BB}$ and $\varepsilon_{AB}$ would then represent the sums of the interaction energies between pairs $A$-$A$, $B$-$B$ and $A$-$B$. Fundamentally the formula would only change if we considered the interaction between e.g. sets of three or four atoms.

If $\varepsilon_{AB}$ is more negative than the mean of $\varepsilon_{AA}$ and $\varepsilon_{BB}$, then $\Delta U$ is negative. In that case there is a stronger attraction between the atoms $A$ and $B$ than between atoms of the same type. During the mixing, therefore, heat is liberated. The entropy effect is here supported by the energy effect and the homogeneous solution, just as in the previous case, is the thermodynamically stable condition. Contrary to the previous case, however, there here occurs (at a suitable atomic ratio) a tendency towards *ordering* in the sense that each $A$ atom is preferentially surrounded by $B$ atoms and each $B$ atom by $A$ atoms At temperatures such that the absolute value of $\varepsilon_{AB} - \frac{1}{2}(\varepsilon_{AA} + \varepsilon_{BB})$ is large compared to $kT$, the atoms can submit to this ordering provided that a perceptible diffusion can still occur at this temperature.

A typical example of an alloy in which ordering occurs, is $\beta$-brass. This phase occurs in the system copper-zinc. The stability range extends from about 40 to 50% zinc. The system has a body-centred cubic structure. At high temperatures the entropy-effect prevails, so that the copper and zinc atoms are distributed at random among the lattice points. At lower temperatures (below approx. 450 °C) the energy effect predominates, so that ordering occurs. Roughly speaking, each Zn-atom surrounds itself by eight Cu-atoms and each Cu-atom by eight Zn-atoms (*fig. 3*). This ordered structure is designated $\beta'$-brass.

If $\varepsilon_{AB}$ is less negative than the mean of $\varepsilon_{AA}$ and $\varepsilon_{BB}$, then $\Delta U$ is positive. The attraction between identical atoms is then stronger than the attraction between different atoms. During isothermal mixing, the internal energy increases, i.e. heat is absorbed. The entropy effect promotes the mixing, but the energy effect promotes separation. At low temperatures the latter effect prevails and the system consists of two phases: virtually pure $A$ and virtually pure $B$. At rising temperature, however, the entropy effect causes increasing mutual solubility. For a system satisfying the conditions of the previous section the mutual solubility as a function of the temperature can be evaluated as follows.

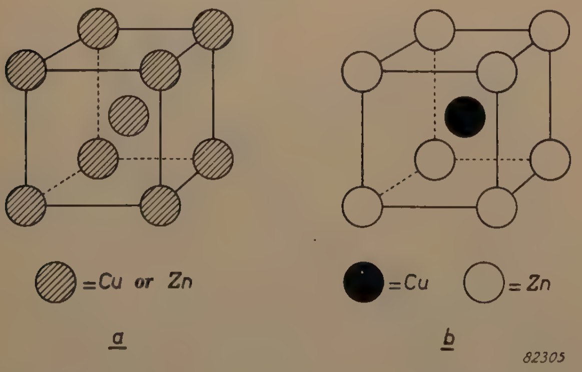


Fig. 3. The unit cell of brass with the composition 50% Cu + 50% Zn.
*a*) At high temperature ($\beta$-brass), the atoms are distributed at random among the lattice points.
*b*) At lower temperatures ($\beta'$-brass), the atoms are ordered, as this is more favourable from the point of view of the internal energy.

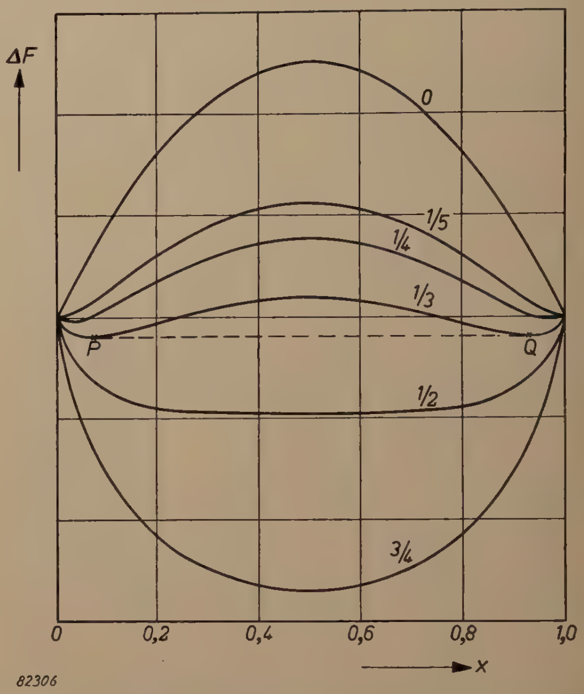


Fig. 4. Free energy of mixing $\Delta F$, according to (III, 20) and (III, 21), as a function of the atomic fraction $x$ for various values of $kT/C$. At low temperatures, segregation occurs as this is more favourable from the point of view of internal energy. For a given temperature, the compositions of the two co-existing phases are defined by the points of contact $P$ and $Q$ of the (dashed) double tangent to the $\Delta F$-curve.

The increase in the free energy due to homogeneous mixing is given, according to (III, 16) and (III, 19) by

$$\Delta F = NCx\,(1-x) + NkT\,\{x \ln x + (1-x) \ln (1-x)\}, \qquad \text{(III, 20)}$$

where

$$C = z\left\{\varepsilon_{AB} - \frac{\varepsilon_{AA} + \varepsilon_{BB}}{2}\right\} \;.\;.\;. \quad \text{(III, 21)}$$

*Fig. 4* shows $\Delta F$ as a function of $x$ for various values of the ratio $kT/C$. The curve for $T = 0$ ($kT/C = 0$) is obviously identical to the $\Delta U$-curve[5]).

---

[5]) It will be clear that the expressions (III, 20), (III, 21) are only justified for small values of $C/kT$. In fact, if $C$ differs from zero, some ordering will immediately occur, causing the entropy of mixing to become smaller than that relating to a random distribution; the energy of mixing will also change.

Throughout a wide temperature range the $\Delta F$-curves show two minima, which approach each other as the temperature rises. The $\Delta F$-curves are symmetrical with respect to the vertical axis $x = 0.5$, which is evidently due to the symmetrical nature of $x$ and $(1-x)$ in (III, 20). The double tangent $PQ$ to a $\Delta F$-curve is thus parallel to the abscissa. The free energy has a minimum in the state of equilibrium, i.e. the homogeneous mixtures (solutions) represented by the branches of the curve to the left of $P$ and to the right of $Q$, are stable. The portion of the $\Delta F$-curves situated between $P$ and $Q$, on the other hand, represent non-stable solutions. By a separation into two phases with compositions $P$ and $Q$, the free energy of a homogeneous alloy can be lowered to a point on the straight line $PQ$. The points $P$ and $Q$, therefore, indicate the solubility values sought. The position of these points can be found by putting the differential coefficient of (III, 20) equal to zero:

$$\frac{\mathrm{d}(\Delta F)}{\mathrm{d}x} = NC\,(1-2x) + NkT\,\{\ln x - \ln\,(1-x)\} = 0, \qquad \text{(III, 22)}$$

from which it follows:

$$\frac{x}{1-x} = \mathrm{e}^{-C(1-2x)/kT}. \quad \ldots \text{(III, 23)}$$

For the construction of the curve representing the solubility as a function of the temperature it is more convenient to put the relationship between $x$ and $T$ in the form

$$T = \frac{C}{k}\,\frac{(1-2x)}{\ln\{(1-x)/x\}}. \quad \ldots \text{(III, 24)}$$

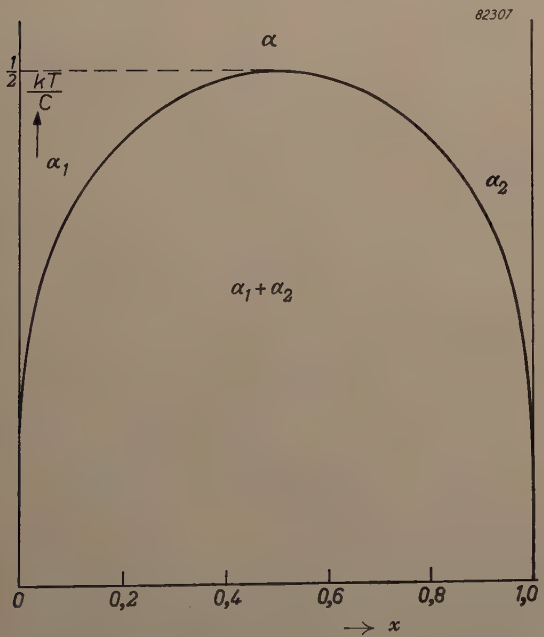


Fig. 5. Solubility curve (segregation curve) according to (III, 23) or (III, 24), giving the relationship between $kT/C$ and the composition ($P$ and $Q$ in fig. 4) of the co-existing phases $\alpha_1$, $\alpha_2$.

*Fig. 5* gives the solubility curve according to (III, 23) or (III, 24). The diagram shows that above a certain temperature complete miscibility occurs.

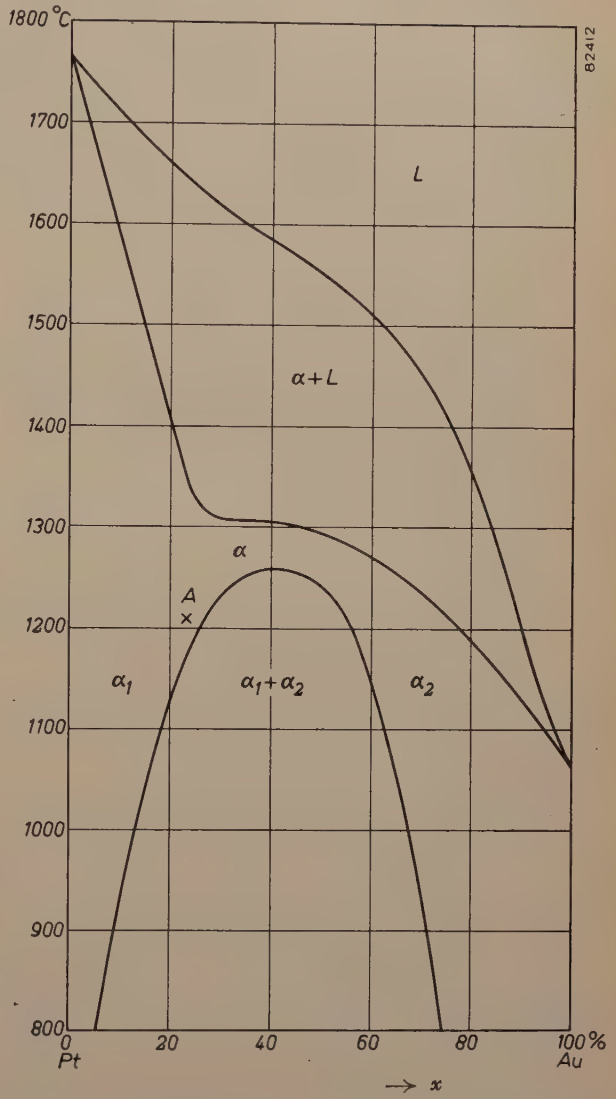


Fig. 6. Phase diagram of the system platinum-gold, which, even in the solid state, shows a complete curve of the type of fig. 5. $\alpha$, $\alpha_1$ and $\alpha_2$ indicate solid phases; the liquid phase is denoted by $L$.

This "critical" temperature can be evaluated as follows. A $\Delta F$-curve with two minima will also have two points of inflection, at which we have

$$\frac{d^2(\Delta F)}{\mathrm{d}x^2} = -2\,NC + NkT\left(\frac{1}{x} + \frac{1}{1-x}\right) = 0,$$

whence
$$2C = kT\left(\frac{1}{x} + \frac{1}{1-x}\right). \quad \ldots \text{(III, 25)}$$

At the temperature rises, the points of inflection approach each other to coincide ultimately at $x = {}^1/_2$, when the critical temperature $T_k$ is reached. From (III, 25) it follows:

$$T_k = \frac{C}{2k}. \quad \ldots\ldots \text{(III, 26)}$$

Actual solubility curves never show the symmetrical form of fig. 5. This is mainly due to the fact that the assumption made above, viz. that the energy may be taken as the sum of interactions between pairs only, involves a certain approximation [4]). A complete solubility curve in the sense described above occurs, e.g., in the system platinum-gold (*fig. 6*). In most cases of binary systems with a positive $\Delta U$-value, however, melting phenomena occur long before the critical mixing temperature is reached. An example of this is the system silver-copper (*fig. 7*).

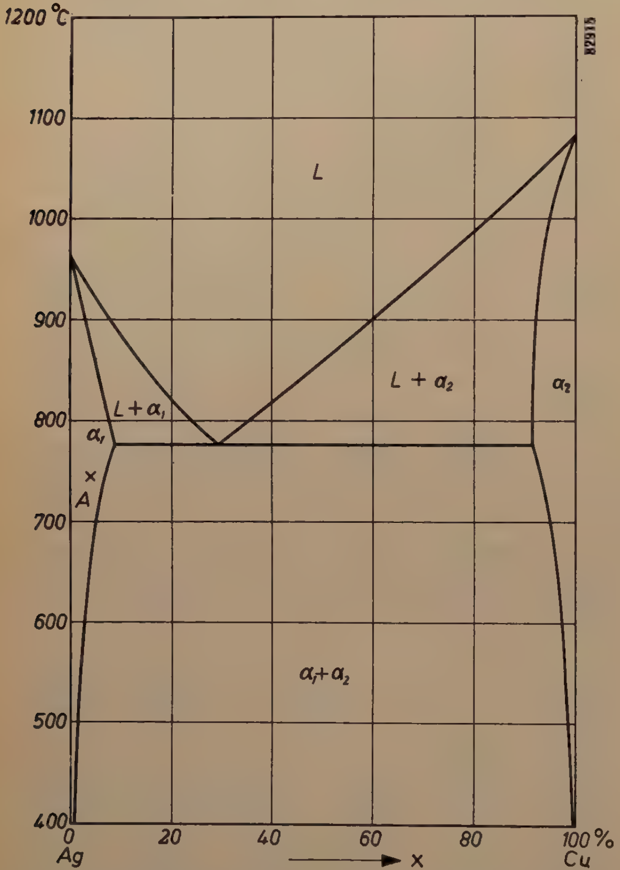


Fig. 7. Phase diagram of the system silver-copper. In this case melting phenomena occur long before the critical mixing temperature is reached.

The decreasing solubility at decreasing temperature occurring in systems of the type discussed here, is often utilized industrially, where it is known as precipitation hardening. This is done by rapidly cooling an alloy from the homogeneous region (e.g. from point $A$ in figs 6 and 7) down to a temperature at which there are two phases in the state of equilibrium and at which virtually no diffusion occurs. Owing to the rapid cooling the thermodynamically required separation cannot occur and a super-saturated homogeneous solution is obtained. By subsequently heating the alloy at a suitable, not excessively high temperature, incipient precipitation occurs, which is accompanied by a substantial increase in the hardness.

## Rubber elasticity

Rubber and rubber-like materials belong to the class of substances known as high polymers. Their molecules take the form of long chains, the links of which consists of groups of relatively few atoms (monomers). There may be several hundreds of identical monomers in the molecule. These monomeric groups are bound together by real chemical bonds. The atoms in these molecules nevertheless have a certain mobility with respect to each other in the sense that one bond can rotate around the other at a constant angle. In *fig. 8* this is shown schematically for a simple case, the black spots representing atoms of the chain. Owing, to this rotational freedom, the chain molecules can assume an enormous number of different forms. Against the one chance of the molecule occurring in the form of a stretched chain there exist countless other possibilities of each molecule occurring in a coiled or tangled form (see *fig. 9*).

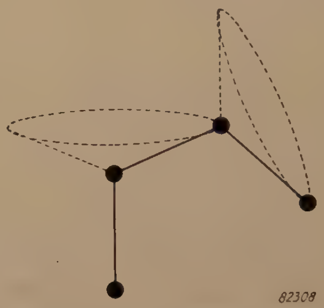


Fig. 8. Macro-molecules often possess a great flexibility owing to the fact that the chemical bonds, though of constant valence angle, can occupy different positions with respect to each other.

The entropy of a high-polymer substance will be greater when the chain molecules have an irregularly coiled form than when they are stretched out and

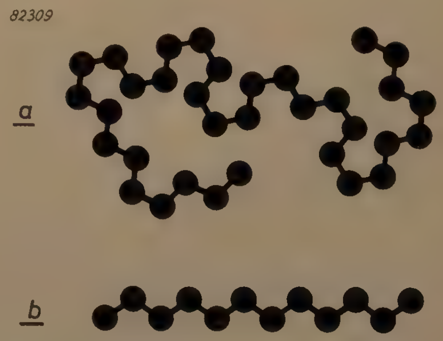


Fig. 9. Part of a molecular chain: *a*) coiled at random, *b*) stretched. For simplicity case (*a*) is depicted in a flat plane (the same applies to fig.10).

lie in parallel arrangement: this follows from the relationship $S = k \ln m$, discussed in I ($m$ = number of micro-states). In some polymers, e.g. cellulose, the latter condition is so much more favourable energetically than the former, that the tendency towards a maximum value of the entropy is overcome by the tendency towards a minimum value of the energy, so that the ordered (stretched) condition prevails. In rubber the difference in energy between the two conditions is only very slight, so that the entropy effect prevails and the molecules have a coiled form. By exerting a tensile stress in a rubber rod, the chains can be stretched, during which considerable elongation of the whole rod occurs. After the stress is removed, the thermal motion of the links of the chain restore the chains into the coiled and thus shortened form. The fact that also the rod as a whole returns to its initial form is explained by the presence of cross-links between the polymer chains. These links are established during the vulcanizing process of the rubber, which gives rise to a wide-meshed network (*fig. 10*). In normal cases about 1% of the monomeric groups are linked by cross-links in this process.

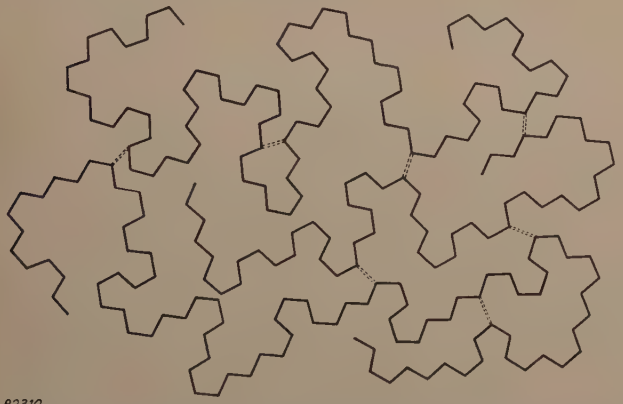


Fig. 10. Vulcanization of rubber has the effect that the long rubber molecules (——) are bound together by short cross-links (=‒=). These cross-links may consist e.g. of two sulphur atoms: —S—S—.

Some of the remarkable thermal properties of rubber can be explained by the foregoing, e.g. the negative coefficient of expansion of strongly stretched rubber. The contraction at rising temperature and constant load is caused by the increasing thermal agitation of the links of the chain molecules, i.e. the increased tendency of the molecules to assume their shortened, coiled form. The parallel alignment of the molecules of natural rubber under considerable stress is a kind of crystallization phenomena. The crystalline nature of the rubber in this condition can be established by X-ray diffraction. After removal of the load the crystalline parts "melt" again.

For a thermodynamical treatment of the subject we start from the first law

$$dU = dQ + dW, \quad . \quad . \quad . \quad . \quad \text{(III, 27)}$$

where $dQ$ represents the heat supplied to the system and $dW$ the work exerted upon it. Suppose that the system is a rubber rod of unit diameter and length $l$, and it is stretched to a length $l + dl$. If we disregard the very small amount of energy spent to bring about the small volume change, $dW$ is given by:

$$dW = \sigma\, dl,$$

where $\sigma$ is the applied stress. If the stretching takes place reversibly, $dQ = TdS$ (if $S$ is the entropy of the rubber rod); $\sigma$ is the equilibrium value of the stress in that case. Equation (III, 27) can now be written as:

$$dU = TdS + \sigma\, dl \quad . \quad . \quad . \quad . \quad \text{(III, 28)}$$

At constant temperature, therefore, we have

$$d(U - TS) = dF = \sigma\, dl,$$

and hence

$$\sigma = \frac{\partial F}{\partial l} = \frac{\partial U}{\partial l} - T\frac{\partial S}{\partial l}. \quad . \quad . \quad \text{(III, 29)}$$

For many types of rubber it has been found experimentally that in the case of a large, constant elongation $\Delta l$, $\sigma$ is approximately proportional to the absolute temperature $T$. This confirms the statement made above that $\partial U/\partial l$ (the increase of the internal energy with the elongation and therefore with the uncoiling of the polymer chain) is very small for this class of substances. This means the the elastic properties of rubberlike substances are a result merely of the decreasing entropy, during stretching and not on the accompanying small change in the energy. For the majority of solids, e.g. metals, the reverse is true. These materials become cool slightly when adiabatically and elastically stretched, whereas rubber heats up slightly under the same conditions. This rise in temperature can be partly explained by the fact that the total entropy does not change with a reversible adiabatic process, because $dS = dQ_{\text{rev}}/T$. The decrease in the entropy, corresponding to the increasing order during strain, is therefore compensated by an increase in the vibrational entropy of the molecules; this is accompanied by an increase in temperature.

The behaviour of a "perfect" rubber, for which $\partial U/\partial l = 0$ and whose elasticity may be regarded as a pure entropy effect, is completely analogous to that of a perfect gas. For a reversible volume change of a perfect gas,

$$dU = TdS - pdV,$$

so that at constant temperature

$$d(U-TS)=dF=-p\,dV,$$

$$p=-\frac{\partial F}{\partial V}=-\frac{\partial U}{\partial V}+T\frac{\partial S}{\partial V}.$$

At constant volume, the pressure of a gas is found to be approximately proportional to the absolute temperature, i.e. the internal energy $U$ of gases is nearly independent of the volume. If the gas is "perfect" $\partial U/\partial V=0$. The pressure of such a gas can thus also be described as an entropy effect.

## Solutions of polymers

We have seen that the entropy of mixing of a homogeneous solid alloy, so far as it depends upon the configuration of the atoms concerned, is given by formula (III, 16). Strictly speaking this formula is only applicable when the *energy* of mixing $\Delta U=0$. If $\Delta U$ differs from zero, the various configurations no longer possess the same probability. In this case the formula provides a sufficient approximation only for temperatures at which $kT$ is considerably greater than $\varepsilon_{AB}-\frac{1}{2}(\varepsilon_{AA}+\varepsilon_{BB})$.

No reference was made above to any difference in size between the atoms or molecules of the mixture. This was not necessary for the *solid* alloys under consideration, in view of the fact that two metals with greatly different atomic radii will not show any appreciable miscibility in the solid state.

*Liquid* homogeneous mixtures, however, often consist of molecules (or atoms) of widely differing sizes. Solutions of polymers are extreme cases of such mixtures. In the following we shall derive an approximate expression for the entropy of mixing of solutions of this kind.

Let us first consider a liquid mixture of molecules $A$ and $B$, which have the same shape and size and furthermore show no preference for similar or dissimilar neighbouring molecules ($\Delta U=0$). In such a mixture, a molecule $B$ may be substituted for a molecule $A$ without any effect on the immediate surroundings. As is known from research on liquids, the immediate surroundings differ little from those in the crystalline state. The arrangement is merely slightly less ordered, so that the order does not extend over such a long range as in the crystalline form. Since we are mainly interested in the bonds between neighbours, it is permissible to use a quasi-crystalline model (*fig. 11*) to represent liquids of the type under consideration. Formula (III, 16) is then applicable without modification. This formula assumes the interchangeability of the molecules $A$ and $B$. If however the molecules differ considerably in size or shape, then the entropy of mixing, even when the energy of mixing $\Delta U$ is zero, is no longer given by (III, 16). For example, a direct interchangeability of molecules of $A$ and $B$ is out of the question if the molecules $A$ are spherical and occupy one site in the diagram of fig. 11 and the molecules $B$ are dumb-bell shaped and occupy two sites (*fig. 12*).

The extreme case of polymers (macro-molecules) dissolved in a micro-molecular solvent can be idealized as follows [6]). It is assumed that a polymer behaves as if it were divided into a large number of mobile segments, each having the same size as one molecule of the solvent. It is further assumed that each segment occupies one site in the quasi-lattice and that the adjacent segments of a chain occupy adjacent sites (*fig. 13*). If the solution contains $N_1$ molecules of solvent and $N_2$ polymer molecules, each consisting of $p$ segments, then there are $(N_1+pN_2)$ sites in the sense of figs 11-13. The fractions

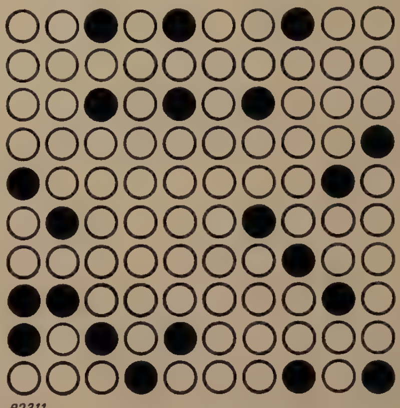


Fig. 11. Schematic representation of a solution of molecules $B$ (black) in a liquid of molecules $A$ (white). The molecules $A$ and $B$ are of about the same size and are statistically distributed in the quasi-lattice.

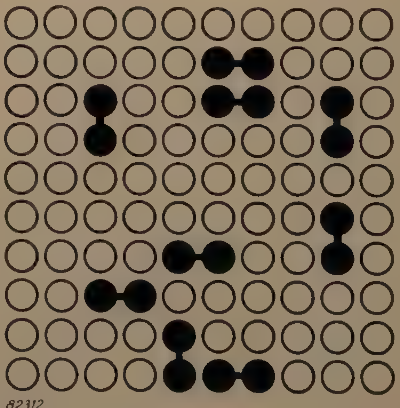


Fig. 12. Schematic representation of a statistical distribution of dumb-bell shaped molecules $B$ (black) in a liquid of spherical molecules $A$ (white). The dumb-bell shaped molecules occupy two adjacent sites of the quasi-lattice.

---

[6]) Cf. e.g. P. J. Flory, Principles of polymer chemistry, Cornell University Press, New York 1953.

of the total volume occupied by both components are given by

$$\varphi_1 = \frac{N_1}{N_1 + pN_2}, \qquad \varphi_2 = \frac{pN_2}{N_1 + pN_2}. \qquad \text{(III, 30)}$$

An expression for the entropy of mixing, in which $\varphi_1$ and $\varphi_2$ are to play an important part, is now derived as follows. First consider a state in which all $(N_1 + pN_2)$ sites are unoccupied. The first segment of the first polymer molecule can thus be accommodated in $(N_1 + pN_2)$ different ways. Once a certain site hasb een decided upon, the second segment of the same polymer molecule can be accommodated in the quasi-lattice in $z$ different ways, $z$ being the number of adjacent sites of any given site in the lattice (the so-called *co-ordination number*).

Each of the subsequent segments (the 3rd, 4th, . . ., $p$th of the same polymer-molecule has $(z-1)$ sites are available, because one of the $z$ adjacent sites is already occupied by the first segment. The first polymer molecule can, therefore, be accommodated in

$$(N_1 + pN_2)\, z\, (z-1)^{p-2} \quad . \; . \; . \quad \text{(III, 31)}$$

different ways [7]).

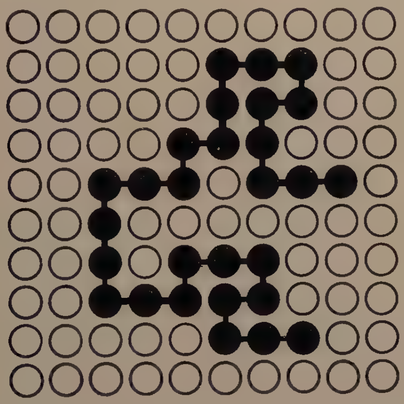


Fig. 13. Schematic representation of an arbitrarily coiled polymer molecule (black) in a solvent of the same type as that in figs. 11 and 12.

Let us now assume that the quasi-lattice already accommodates $(k-1)$ polymer molecules. In how many different ways can the $k$th polymer molecule now be accommodated? The first segment of this molecule can be placed in any one of

$$N_1 + pN_2 - (k-1)\, p = N_1 + p\, (N_2 - k + 1)$$

different sites. The second segment of the $k$th polymer molecule, unlike that of the first, cannot be accommodated in $z$ different ways, since there is a possibility that one of the adjacent sites is occupied by a segment of one of the $(k-1)$ polymer molecules already accommodated. This possibility can be accounted for with a reasonable accuracy by multiplying $z$ by the ratio of the number of sites still unoccupied to the total number of sites

$$\frac{N_1 + p(N_2 - k + 1)}{N_1 + pN_2}.$$

A similar reduction factor can be applied to each factor $(z-1)$ for the number of ways in which the 3rd, 4th, ... $p$th segment can be accommodated. The $k$th polymer molecule can, therefore, be placed in the lattice in

$$\frac{\}N_1 + p(N_2 - k + 1)\{^p\, z\, (z-1)^{p-2}}{(N_1 + pN_2)^{p-1}} \qquad \text{(III, 32)}$$

different ways. For $k = 1$ this expression reduces to the form (III, 31). The total number of ways in which $N_2$ identical polymer molecules can be accommodated in the lattice is given by the product of the $N_2$ terms (III, 32) which are obtained by subsequently substituting for $k$ the values 1, 2...., $N_2$ Since in our case the polymer molecules are actually identical, so that no new situation is created when two of them are interchanged, the resulting number has to be divided by $N_2!$. The total number of possible different arrangements of $N_2$ identical polymer molecules among the $(N_1 + pN_2)$ sites of our model is therefore

$$m = \frac{\}z\,(z-1)^{p-2}\{^{N_2}}{N_2!} \times$$

$$\frac{(N_1 + pN_2)^p[(N_1 + p(N_2 - 1)]^p \ldots (N_1 + p)^p}{(N_1 + pN_2)^{(p-1)N_2}}. \; \text{(III, 33)}$$

In each of these arrangements a number of sites $(N_1)$ remain unoccupied. Since in each case these sites can be occupied by molecules of the solvent in one way only, the addition of these molecules does not add to the number $m$. In other words, the expression (III, 33) also gives the number of possible configurations $m_{12}$ of the mixture of $N_1$ solvent molecules and $N_2$ polymer molecules. After rearranging, we obtain the expression:

$$m_{12} = \frac{\}z(z-1)^{p-2}\{^{N_2}\, p^{N_2}}{N_2!\left(\dfrac{N_1}{p} + N_2\right)^{(p-1)N_2}} \; \frac{\left\{\left(\dfrac{N_1}{p} + N_2\right)!\right\}^p}{\left\{\left(\dfrac{N_1}{p}\right)!\right\}^p}. \quad \text{(III, 34)}$$

The entropy of mixing is given by

$$\Delta S = k \ln m_{12} - k \ln m_1 - k \ln m_2, \quad \text{(III, 35)}$$

where $m_1$ is the number of possible arrangements of $N_1$ solvent molecules in the solvent itself ($N_1$ sites), and $m_2$ for the number of possible arrangements of the $N_2$ polymer molecules in the polymer itself ($pN_2$ sites).

---

[7]) Here certain factors have been overlooked. We shall return to these points at the end of this article.

It is clear that $m_1 = 1$ and hence $k \ln m_1 = 0$. For $m_2$ this does not apply because of the many possible coiled configurations of the polymer chains. If it is assumed that the polymer molecules in the undiluted state behave in an analogous way as in the solution, then $m_2$ is found by putting $N_1 = 0$ in (III, 34):

$$m_2 = \frac{\left\{z(z-1)^{p-2}\right\}^{N_2} p^{N_2} (N_2!)^p}{N_2!\, N_2^{(p-1)N_2}}. \quad \text{(III, 36)}$$

From the last three formulae it follows that

$$\Delta S = k \ln \left[ \frac{\left\{\left(\frac{N_1}{p} + N_2\right)!\right\}^p}{\left\{\left(\frac{N_1}{p}\right)!\right\}^p \left(N_2!\right)^p} \frac{N_2^{(p-1)N_2}}{\left(\frac{N_1}{p} + N_2\right)^{(p-1)N_2}} \right].$$

Employing Stirling's formula in the approximation (I, 3), we find:

$$\Delta S = -k N_1 \ln \frac{N_1}{N_1 + pN_2} - k N_2 \ln \frac{pN_2}{N_1 + pN_2}. \quad \text{(III, 37)}$$

If in the previous evaluation we had erroneously considered the polymer molecules as being equivalent to the micromolecules, then the earlier formula (III, 16) would again have been obtained, which in the same notation as (III, 37) has the form:

$$\Delta S = -k N_1 \ln \frac{N_1}{N_1 + N_2} - k N_2 \ln \frac{N_2}{N_1 + N_2}. \quad \text{(III, 38)}$$

This is obtained from (III, 37) by putting $p = 1$.

If, on the other hand, the $pN_2$ segments of the polymer molecules had been considered as separate molecules, then we would have obtained the relationship

$$\Delta S = -k N_1 \ln \frac{N_1}{N_1 + pN_2} - kpN_2 \ln \frac{pN_2}{N_1 + pN_2}. \quad \text{(III, 39)}$$

The conclusions from (III, 37) will be intermediate between those of (III, 38) and (III, 39).

Using equation (III, 30), we may also write (III, 37) as:

$$\Delta S = -kN_1 \ln \varphi_1 - kN_2 \ln \varphi_2, \quad . \ . \ \text{(III, 40)}$$

where $\varphi_2 = 1 - \varphi_1$.

This formula has a similar form to that of the earlier formula (III, 16) for the entropy of mixing; instead of the mole fractions $x_1$ and $x_2 = 1 - x_1$, however, here the volume fractions $\varphi_1$ and $\varphi_2 = 1 - \varphi_1$ occur.

From the derivation of formula (III, 37) it appears that any symmetry factor may be disregarded in (III, 31), since this factor would occur both in the expression for $m_{12}$ and in that for $m_2$, so that it would be cancelled out in the expression for $\Delta S$ (cf. formula (III, 35)). For the same reason we may overlook the fact that for the 3rd, 4th ..., $p$th segments of a polymer molecule less than $(z - 1)$ sites are available, because a polymer molecule is not as a rule perfectly flexible and, moreover, one or more of the $(z - 1)$ sites may be occupied by previous segments.

Several thermodynamical properties of solutions of polymers, such as osmotic pressure and miscibility can be understood by means of formulae( III,37), which also gives us a picture of the difference between this type of solution and that of common micro-molecular solutions.

# IGNITION OF "PHOTOFLUX" FLASH-BULBS WITH THE AID OF A CAPACITOR

by J. A. de VRIEND.

771.448.4:621.319.4

*In photography by means of "Photoflux" flash-bulbs, the "time to peak" or "time lag", is dependent on the total resistance of the circuit. The use of a higher voltage appreciably reduces this time lag, and the high voltage can be obtained from a capacitor in accordance with a method that is already widely used.*

## Ignition time, time to peak and circuit resistance

A large number of photographs are nowadays taken with the aid of flash-bulbs such as the "Photoflux".

The simplest form of ignition is obtained from a dry battery with switch (*fig. 1*) [1]), the latter being usually an electrical contact incorporated in the camera and synchronized with the shutter release.

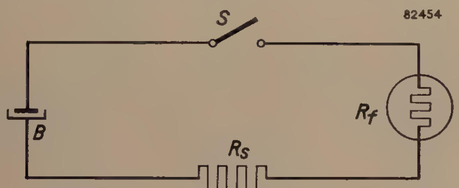


Fig. 1. Ignition of "Photoflux" flash bulb by means of a battery ($B$). $R_f$ = resistance of igniter wire in the bulb, $R_s$ = external resistance in series with bulb, $S$ = switch.

The maximum emission of light does not occur immediately the circuit is closed; there is even a certain delay before the bulb gives any light at all, but subsequently the intensity rises very steeply (*fig. 2*). The time $t_f$ which precedes the commencement of the light is necessary for the igniter wire in the "Photoflux" bulb to reach the required temperature, after which combustion occurs and light is produced. The time elapsing between the moment of closing the circuit and the flash peak is called the "time to peak" ($t_m$) in fig. 2).

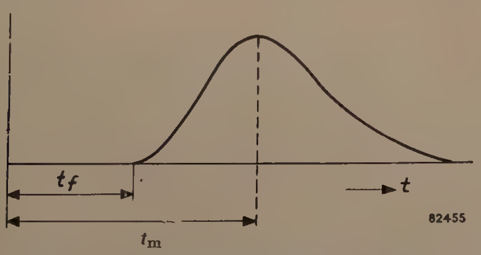


Fig. 2. After the circuit has been closed, a time $t_f$ elapses before the igniter wire reaches the required temperature (heating time), and a time $t_m$ before the flash reaches its maximum intensity (time to peak, or time lag).

Particularly for exposures of less than 1/50 sec, it is required that this time lag shall be roughly constant [2]), viz., according to current standards, 20 milliseconds. To ensure this, the ignition time $t_f$ must not exceed a certain fraction of $t_m$, for example $t_f < 0.2\ t_m$. It is found that, from the moment the igniter wire reaches the firing temperature, "Photoflux" flash bulbs always attain maximum light emission, or flash peak, in roughly the same space of time. This means that, in fig. 2, $t_m - t_f$ is constant.

The requirement that $t_f$ shall not be greater than $0.2\ t_m$ places a lower limit on the amount of power to be supplied to the igniter wire. If the total electrical resistance in the firing circuit is increased for one reason or another, the power delivered to the igniter wire by the battery is of course reduced. This means a longer ignition time $t_f$, possibly even beyond the permissible limit.

The total circuit resistance comprises the resistance of the igniter wire itself, the internal resistance of the battery and any other additional resistance that may be in the circuit. Whereas the resistance of the igniter wire is constant by reason of the manufacturing process, that of the battery increases as the battery runs down and extra resistances may be introduced by dirty switch contacts or poor connections. If the leads are longer than usual, this will also increase the resistance.

If a capacitor with resistor in series is added to the circuit shown in fig. 1, a practice that has been more and more widely adopted in recent years, several improvements are assured, the most important of these being that higher circuit resistances can be tolerated, without exceeding the specified firing time. A simple calculation will illustrate the difference between ignition by means of a battery only and that by means of a battery with capacitor.

---

[1]) For a description of "Photoflux" flash bulbs see Philips tech. Rev. **12**, 185-192, 1950/51 and **15**, 317-321, 1953/54. For a more general review see G. D. Rieck and L. H. Verbeek, "Artificial light and Photography", Philips Technical Library, 1950.

[2]) See first article mentioned in note [1]), p. 191.

### Direct battery ignition

Let us assume that the resistance of the igniter wire in the circuit in fig. 1, ($R_f$) is constant. This is of course not strictly correct, in view of the fairly considerable rise in temperature (and hence also in the resistance), but for a comparison of the two methods of ignition the assumption is permissible and greatly simplifies the calculation.

If $R_i$ is the internal resistance of the battery, $R_s$ the external series resistance, $E_b$ the e.m.f. of the battery, and $P_f$ the power supplied to $R_f$, we have:

$$P_f = R_f\left(\frac{E_b}{R_f+R_i+R_s}\right)^2 . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The question is now: what should be the minimum value of $P_f$ to ensure that the flash bulb ignites within the specified time?

Disregarding the loss of heat to the surroundings during the very short heating time (only a few millisec) the minimum amount of power will be determined by the maximum permissible heating time $t_{f\max}$ and the required heating energy $Q_f$ (product of thermal capacity and ignition temperature of the igniter wire):

$$P_f \geq \frac{Q_f}{t_{f\max}} .$$

Care in manufacture ensures that $Q_f$ is practically constant. With (1) we can therefore write:

$$E_b \geq \left(1+\frac{R_i}{R_f}+\frac{R_s}{R_f}\right)\sqrt{\frac{R_f Q_f}{t_{f\max}}} \quad . \quad . \quad . \quad . \quad (2)$$

With the following values: $R_f = 2$ ohms, $t_{f\max} = 3.5 \times 10^{-3}$ sec and $Q_f = 3.5 \times 10^{-3}$ joules, equation (2) becomes:

$$E_b \geq \tfrac{1}{2}\sqrt{2}\ (2+R_s+R_i) .$$

This relationship between $E_b$, $R_i$ and $R_s$ is shown graphically in *fig. 3*. A battery voltage of, say, 4.5 V, with $R_i = 2$ ohms, gives, from the graph, a maximum permissible series resistance of 2.4 Ω. If the series resistance is any higher the maximum heating time will be exceeded.

The longer the heating time, the more significant the heat that is lost to the surroundings, as this in turn increases $t_f$. In the extreme case, with very high $R_s$ (and/or $R_i$) these losses will be equal to the applied power before the igniter wire reaches the working temperature, with the result that the bulb cannot ignite at all.

For a higher resistance to be permissible, $E_b$ (fig. 3) must be increased and/or $R_i$ reduced. It should be remembered, however, that there is a relationship between the e.m.f., the internal resistance and the volume of the battery. In practice, batteries having an e.m.f. of at most 4.5 V are generally used. With a higher voltage, either the volume of the battery is too great, or the internal resistance too high. Moreover, $R_i$ increases as the battery runs down, so that the permissible amount of extra series resistance becomes less and less in the course of time.

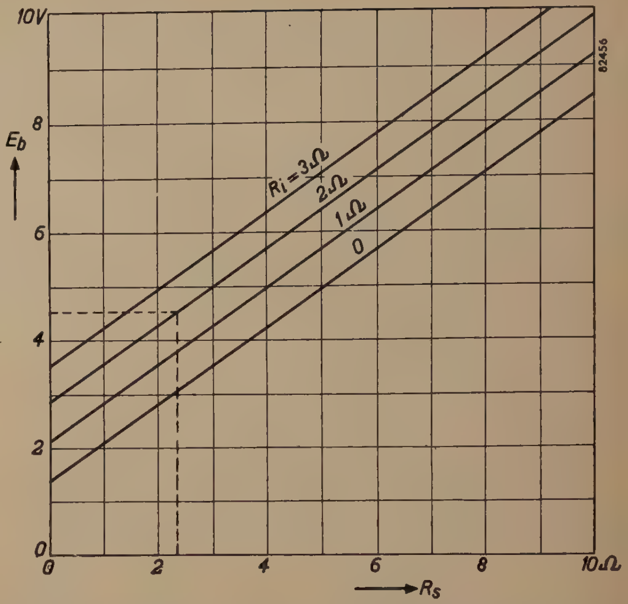


Fig. 3. Calculated values of the maximum permissible series resistance $R_s$ plotted against battery voltage $E_b$ for various battery internal resistances $R_i$, for direct battery ignition.

There are accordingly many objections to direct battery ignition, but these can largely be eliminated by using a capacitor with series resistor.

### Ignition with the aid of a capacitor

The circuit incorporating a capacitor $C$ and series resistor $R_v$ is shown in *fig. 4*.

The internal resistance of the battery does not affect the ignition and therefore no longer causes the power supplied to vary, so that a battery having a high internal resistance can be used, thus making feasible a higher voltage battery, within the limits of a convenient volume.

Another advantage is that the resistance in series with the battery may also be higher, which means such a small load on the battery that its life will be almost as long as with no load at all. The charging current for the capacitor is certainly small, but there should be no objection to this as there is usually sufficient time for recharging (at least a few seconds).

Now, what are the conditions to be imposed on the ignition circuit $R_f$, $R_s$ and $C$?

Let $E_c$ and $E$ denote the values of the capacitor voltage at the commencement and end of the discharge time $t$; then:

$$E = E_c\, e^{-\frac{t}{(R_f+R_s)\,C}}.$$

The initial energy $\frac{1}{2}CE_c^2$ of the charged capacitor is reduced in a time $t$ by an amount

$$Q_c = \tfrac{1}{2}\,C\,(E_c^2 - E^2)\,.$$

This energy is dissipated in the resistors $R_f$ and $R_s$, whereby $R_f$ receives a fraction equal to $R_f/(R_f + R_s)$. Since a heating energy at least equal to $Q_f$ must be delivered in the maximum permissible heating time $t_{fmax}$.

$$\tfrac{1}{2}\,C\,E_c^2\left[1 - e^{-\frac{2t_{fmax}}{(R_f+R_s)C}}\right]\frac{R_f}{R_f + R_s} \geq Q_f \quad . \quad . \quad (3)$$

With some re-arrangement and substitution of the values of $t_{fmax}$, $R_f$ and $Q_f$ we now write:

$$E_c \geq \sqrt{\frac{0.0035\,(2 + R_s)}{C\left(1 - e^{-\frac{0.007}{(2+R_s)C}}\right)}}\,.$$

The equality relation of this expression is shown graphically in *fig. 5*, from which we can at once see the advantage of capacitor ignition in comparison with direct battery ignition. Note, for example that at 22.5 V (a practical value) up to 15.7 Ω in series can tolerated with a capacitor of 130 μF. It is also seen that it is important to use the largest possible capacitance compatible with the dimensions of the capacitor. At 22.5 V and with a 700 μF capacitor the permissible series resistance rises to 27 Ω.

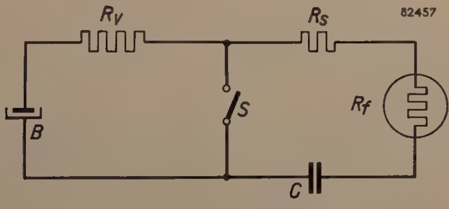


Fig. 4. Ignition with the aid of a capacitor ($C$). This capacitor is charged by a very small current through the resistance $R_f$ of the "Photoflux" bulb and the high series resistance $R_v$. When the circuit is closed, $C$ is discharged across $R_f$. $R_s$ is the extra series resistance present.

In order better to illustrate the curves plotted from equation (3) let us consider two extreme cases.

Writing $z$ as the exponent of e in (3), consider first:

$$z = \frac{2t_{fmax}}{(R_f + R_s)\,C} \ll 1\,. \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

The term in square brackets in equation (3): $F = 1 - e^{-z}$ can be expanded as a series, viz. $z - \frac{1}{2}z^2 + \dots$ For a small value of $z$ the first term of this series is sufficient, and (3) then becomes:

$$E_c^2 \geq \left(1 + \frac{R_s}{R_f}\right)^2 \frac{R_f Q_f}{t_{fmax}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3a)$$

This is the same as equation (2) except that $R_i$ is now missing, which means that a very large capacitor, for which condition (4) is certainly met, can be regarded as a battery without internal resistance. The limiting line $C = \infty$ in fig. 5 is given by equation (3a), and is at the same time the line corresponding to battery ignition when $R_i = 0$.

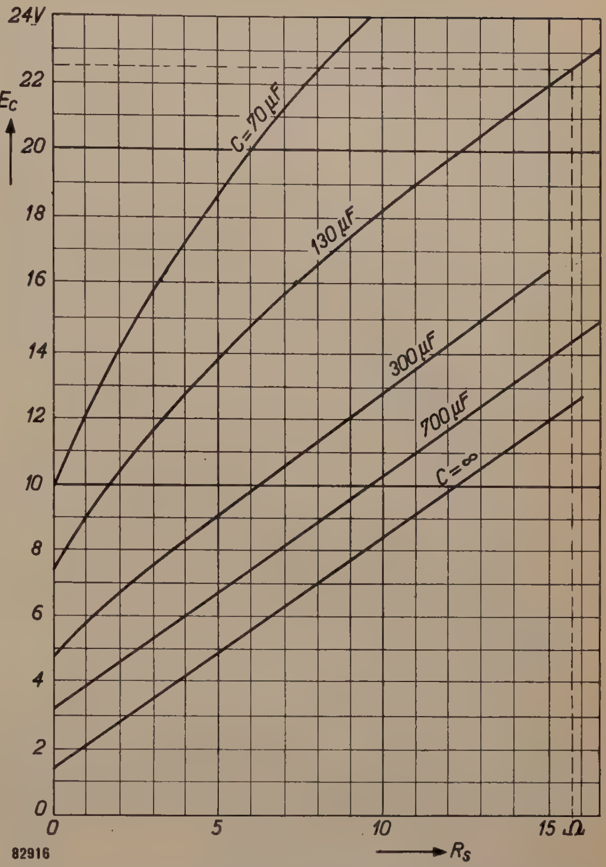


Fig. 5. Calculated values of the maximum permissible resistance $R_s$ in series with the bulb, plotted against the battery voltage $E_c$ for various capacitances $C$, for capacitor ignition.

The second extreme occurs when $e^{-z}$ is very small (capacitor almost wholly discharged). For instance, with

$$z = \frac{2\,t_{fmax}}{(R_f + R_s)\,C} > 5\,. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

$e^{-z}$ is $< 0.01$. Formula (3) then takes the form

$$E_c^2 \geq \frac{2\,Q_f}{C}\left(1 + \frac{R_s}{R_f}\right). \quad . \quad . \quad . \quad . \quad . \quad . \quad (3a)$$

In that part of the graph where (5) applies, the curves are therefore parabolas in which $C$ occurs as parameter. The slope increases with decreasing $C$.

## Practical results

*Fig. 6* illustrates the results of a number of measurements made with direct battery ignition (with an accumulator, i.e. $R_i \approx 0$), and capacitor ignition. For these measurements a known resistance $R_s$ was included in series with the circuit, and the voltage required to raise the igniter wire to the correct temperature in a time $t_{fmax}$ was measured.

The qualitative agreement of figures 3 and 5 with fig. 6 is apparent on examination; quantitatively, the curves approximate closely to quations (2) and (3). The line for direct battery ignition, with $R_i = 0$, is shown in fig. 6 as $C = \infty$, as this extreme case corresponds to that condition. It deviates more from the calculated curve for $C = \infty$ (dotted) as $R_s$ is smaller. This is owing to the fact that the variation in $R_f$ with increase in temperature represents a greater percentage change in the total resistance at small values of $R_s$.

In practice, owing to the inevitable amount of spread, the voltages to be employed will be slightly higher than those measured experimentally. Moreover, as the lines in fig. 6 are not very curved, the relationship between $E_c$, $R_s$ and $C$ can conveniently be represented by straight lines. This gives us *fig. 7*, from which the resistance $R_s$ can be found for given values of $E_c$ and $C$ to yield the specified heating time of 3.5 msec.

In the equally important case of the simultaneous firing of two flash-bulbs, similar calculations are applicable. Capacitor ignition is quite practicable provided that the capacitance is high enough, for the discharge is now divided between two circuits (flash bulbs in series are not to be recommended). It will be seen that in this case the series resistor in each circuit may have to conform to a lower limit as well as an upper one, in order to guarantee a sufficiently uniform distribution of the power. However, a discussion of this point is not within the scope of this article.

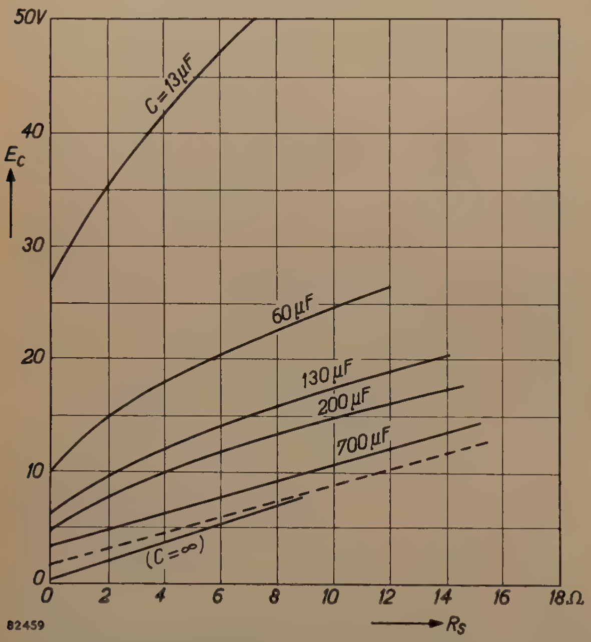


Fig. 6. Maximum permissible series resistance $R_s$ plotted against battery voltage $E_c$ for various capacitances $C$, as obtained experimentally for capacitor ignition. The line denoted by $C = \infty$ refers to direct battery ignition with $R_i = 0$. The broken line is the calculated curve for the latter case.

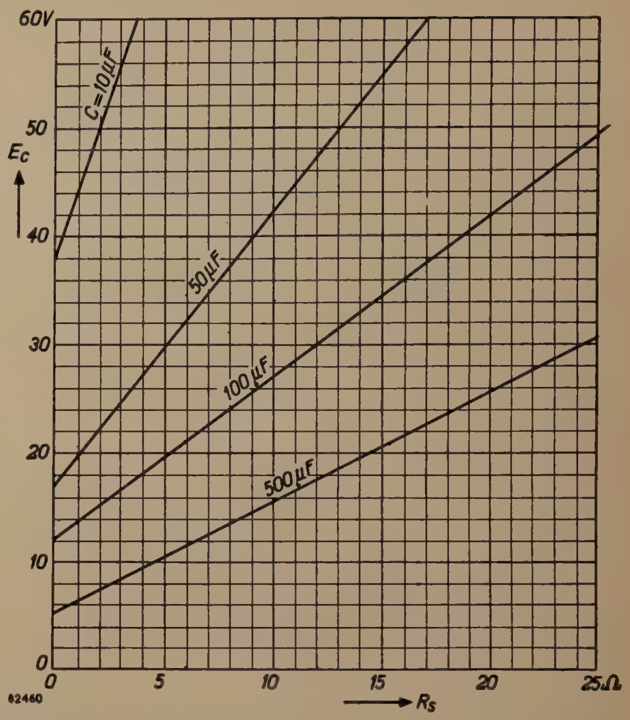


Fig. 7. Capacitor ignition chart for practical use. (Maximum ignition time 3.5 millisec).

**Summary.** When the simplest method of firing a "Photoflux" flash bulb is employed, viz. direct battery ignition, only a small amount of extra resistance in the circuit is permissible if the time lag is to remain sufficiently constant. With higher battery voltages than the customary 4.5 V, the position is better but, as the internal resistance of the battery must be low, again with a view to a constant time lag, the battery would then be too large in size. The well-known method of ignition by means of a capacitor eliminates this difficulty; a voltage of, say 22.5 V can be employed, with a fairly high resistance in series with the bulb. This method is examined quantitatively, and curves are given from which the maximum permissible series resistance for given battery voltage and capacitor value can be read.

# DEMONSTRATION OF THE AUSTENITE-PEARLITE TRANSFORMATION BY MEANS OF THE EMISSION ELECTRON MICROSCOPE

621.385.833:669.112.227.32

In normal electron microscopy the object is irradiated with electrons emitted from a filament. The examination of metals and alloys using a normal electron microscope is therefore largely restricted to pictures of very thin replicas of the metal surface. An important advance, permitting the study of metals at high temperatures, is the development of the emission electron microscope in which the object itself acts as the electron source.

The temperature region in which the observations are to be made is mainly determined by the phenomena to be investigated. Unfortunately, most metal surfaces begin to emit electrons in quantity only at very high temperatures. An adequate emission at relatively low temperatures can be obtained by evaporating onto the surface a very thin layer of electropositive atoms, for example barium or caesium, which lower the work function of the surface of the metal. At the same time, the polycrystalline structure of the metal surface is made visible, because the evaporated atoms are absorbed differently according to the lattice orientations of the various surface crystallites. A change in this structure, for example, on recrystallization, can therefore be directly observed.

An emission electron microscope developed by Philips is shown in *fig. 1*. A number of investigations into recrystallization and phase-transitions in metals and alloys have been carried out with this instrument in recent years [1]). The instrument will be fully described in a later issue of this Review. Here it will only be mentioned that its construction follows as far as possible that of the EM 75 kV electron microscope to be described shortly in these pages [2]). Naturally the emission instrument does not have a filament, nor does it have a condenser lens. Further differences in construction include means by which the object is brought to high potential and the special steps taken in view of the higher vacuum required.

As an example of the type of work which may be done with this emission microscope, some photographs are given illustrating the austenite-pearlite transformation in 0.8 % carbon steel. The metal was prepared under high vacuum to minimize the effect of impurities [3]).

*Fig. 2a* shows the austenitic structure which is stable above 721 °C — a homogeneous solution of carbon in $\gamma$-iron; fig. 2*b* shows the way in which the pearlite grows in the austenite. The pearlite is built up of alternate lamellae of $\alpha$-iron (ferrite) and iron carbide (cementite). The many dark stripes

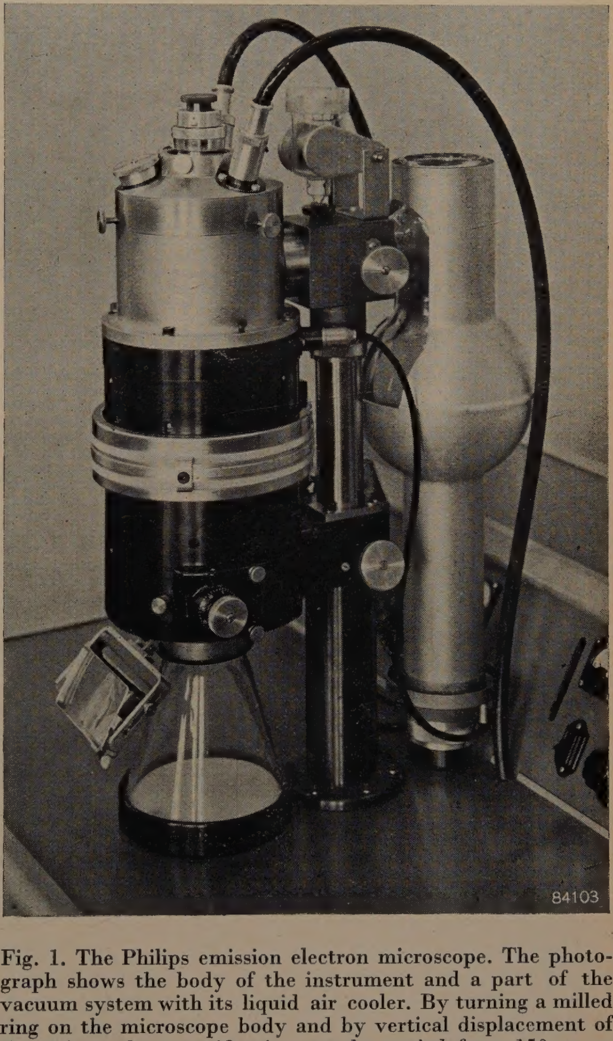


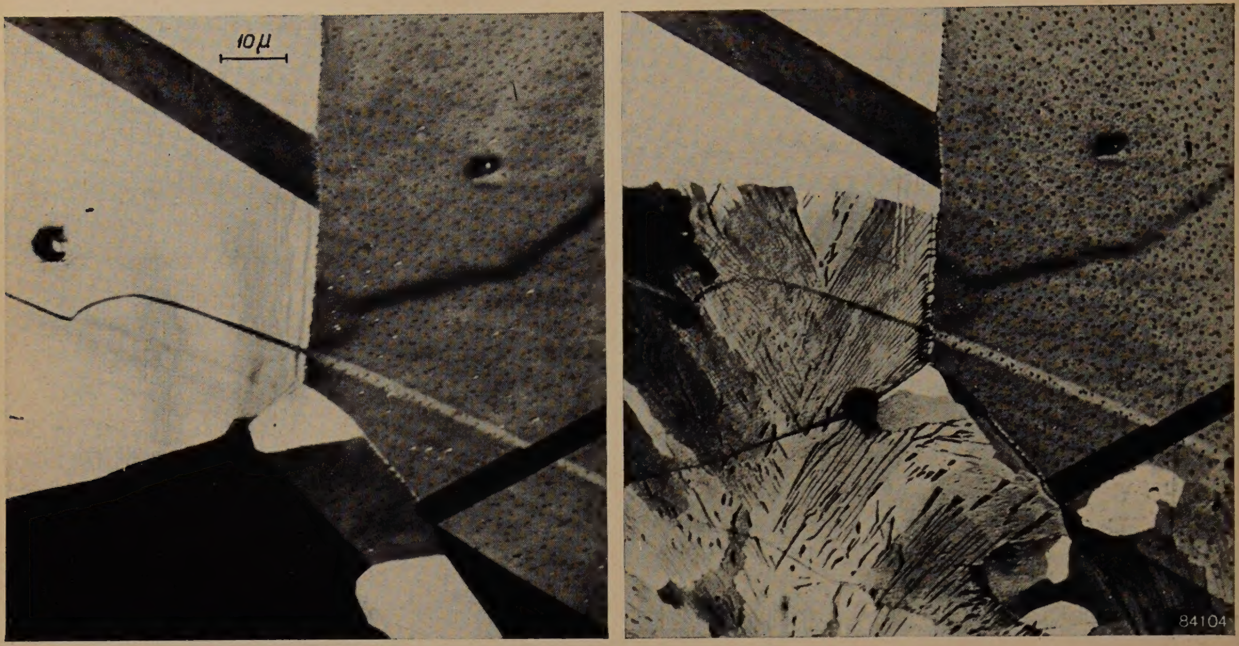
Fig. 1. The Philips emission electron microscope. The photograph shows the body of the instrument and a part of the vacuum system with its liquid air cooler. By turning a milled ring on the microscope body and by vertical displacement of the object, the magnification can be varied from 150 × to 3000 ×. The resolving power is 1000 Å. By means of a built-in camera, a series of 40 pictures can be taken in rapid succession on 35 mm film. If necessary a phenomena can also be filmed by means of a 16 mm cine camera placed under the desk; in this case a transmission fluorescent screen is used and the picture filmed through the glass of the screen.

---

[1]) G. W. Rathenau and G. Baas, Physica **17**, 117-128, 1951.
G. W. Rathenau, L'état solide, 9e Conseil de Physique Solvay, pp. 55-72, Brussels 1952.
G. W. Rathenau and G. Baas, Métaux **29**, 139-150, 1954 (No. 344).
G. W. Rathenau and G. Baas, Acta Met. **2**, 875-883, 1954 (No. 6).

[2]) A. C. van Dorsten and J. B. le Poole, The EM 75 kV, an electron microscope of simplified construction, to appear shortly in this Review.

[3]) J. D. Fast, A. I. Luteijn and E. Overbosch, Preparation and casting of metals and alloys under high vacuum, Philips tech. Rev. **15**, 114-121, 1953/54.

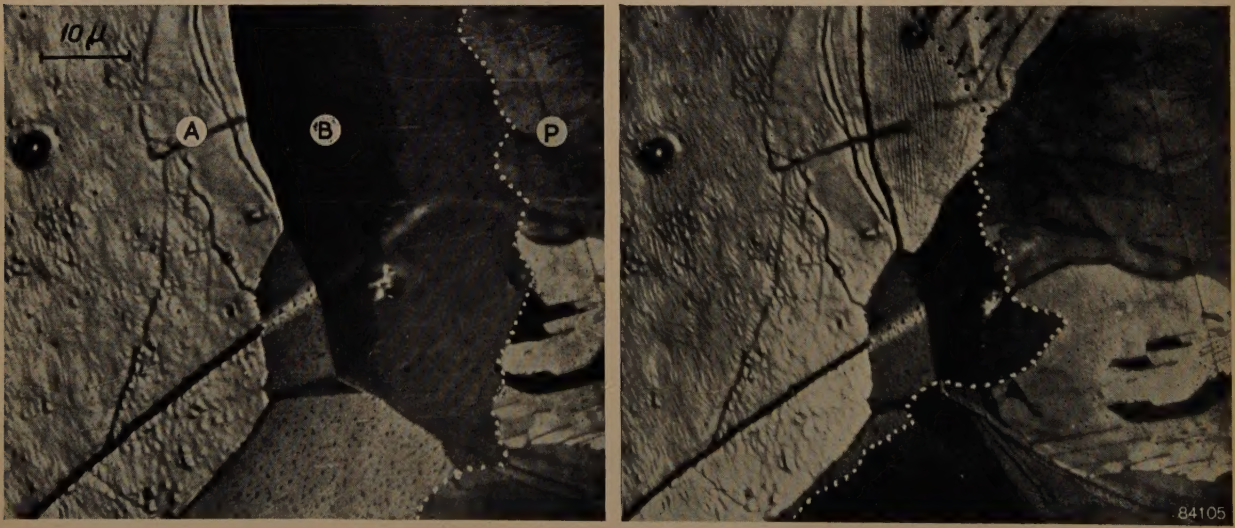


*a* *b*

Fig. 2. Two pictures from a series made during the conversion of austenite into pearlite in 0.81% carbon steel, at a temperature of about 700 °C. The interval between the two pictures is about 2 minutes.
*a*) Austenitic structure
*b*) Austenite partly changed into pearlite.

which can be seen in the pearlite are the cementite lamellae.

*Figs 3a* and *b* show a phenomenon often observed in the austenite-pearlite transformation in this carbon steel: if the pearlite, growing in the austenite matrix, comes in the neighbourhood of a boundary between two austenite crystals, this boundary tends to meet the approaching pearlite. This can be seen in the



*a* *b*

Fig. 3. The boundary between the two austenite crystals *A* and *B* moves in the direction of the approaching pearlite *P*. The dotted line shows the boundary between the pearlite and the austenite.
*a*) Initial situation. *b*) After the boundary has moved.
The presence of the many fine grooves in the part of *B* which has been consumed by *A* shows that the movement of the crystal boundary is a discontinuous process,

photographs at the boundary between the crystals *A* and *B*. The boundary between austenite and pearlite is shown with a dotted line.

In the part of crystal *A* which has grown in this way at the expense of crystal *B*, one can see many fine grooves in the surface, lying very close together (fig. 3*b*). Their presence shows that the displacement of the crystal boundary is a *discontinuous* processs: if the boundary stays still for a moment, such a groove forms as a result of the surface tension at the crystal boundary: the series of grooves shows the successive positions taken up by the crystal boundary.

G. BAAS.

---

# ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of these papers not marked with an asterisk * can be obtained free of charge upon application to the Administration of the Philips Research Laboratory, Eindhoven, Netherlands.

**2140:** K. F. Niessen: Magnetic behaviour of some ferrites (Physica **19**, 1127-1132, 1953).

The variation of the Curie temperature of a series of mixed crystals of nickel ferrite and nickel titanate ($Fe_{2-2a}Ni_{1+a}Ti_aO_4$) with $0 < a < 0.5$, has been measured as a function of the content $a$ of titanium, by E. W. Gorter of this Laboratory. From this data, the distribution of the titanium amongst the tetrahedral (*A*—) and octahedral (*B*—) sites of the spinel is derived on the assumption that nickel ions always occupy *B*-sites. The parameters necessary for this calculation were taken from the form of the Curie temperature curve of nickel zinc ferrite, measured as a function of the nickel content by Guillaud and Roux. The partial magnetizations in $Fe_{2-2a}Ni_{1+a}Ti_aO_4$ were taken parallel and antiparallel. In the case $a \ll 1$ the titanium appears to be distributed statistically (i.e. in the same ratio 1 : 2 as the numbers of *A*— and *B*— sites), but at an increasing content of titanium this metal appears to have an increasing preference for the *A*-sites.

**2141:** J. M. Stevels: Analyse statistique graphique du verre, II (Verres et Réfractaires 7, 281-286, 1953).

The author shows that for glasses of the system $Na_2O$-$CaO$-$SiO_2$ there is a linear relation between the specific refraction $S = [(n_D^2-1)/(n_D^2+2)]/\varrho$ and the dispersion $D = n_F - n_C$, such that $S =$ constant $\times D + f(R)$. The function $f(R)$ is a function of $R$ only (ratio of the number of oxygen ions to the number of $Si^{4+}$ions) and is independent of the content of $Na^+$ and $Ca^{++}$ ions. If qualitative analysis shows that the glass contains only the above-mentioned ions, then the formula makes it possible, by a simple measurement of $S$ and $D$, to determine the numbers of bridging and non-bridging oxygen ions. In an appendix the author discusses the theoretical basis of the formula.

**2142:** H. C. Hamaker: "Average confidence" limits for binomial probabilities (Rev. Int. Statist. Inst. **21**, 17-27, 1953).

A detailed investigation of the "confidence limits" of binomial probabilities as generally used, shows that these limits are fixed in such a way that in the most unfavourable circumstances the percentage of correct decisions is never below the confidence level. In industrial practice it should not be assumed that these unfavourable conditions are always present. Thus it is reasonable to modify the limits in such a way that the percentage of correct decisions is on average equal to the confidence level. The "average confidence" limits introduced in the present paper are designed towards this end. They determine a region of "average confidence" which is narrower than that generally used.

**2143:** F. van Tongerloo: Magnetic and dielectric elements for computers (T. Ned. Radiogenootsch. **18**, 265-285, 1953; in Dutch).

A review of the applications of magnetic and dielectric materials with square hysteresis loops. The first part of the article deals with circuits which depend on a high value of the ratio of remanance to saturation; the second part deals with circuits in which the squareness ratio is important. The article contains no original developments.

**2144*:** E. J. W. Verwey: New developments in synthetic ceramics (Proc. Int. Symp. on reactivity of solids, Gothenburg 1952, Elanders Bocktryckeri A.B., Göteborg, 1953, pp. 703-715).

Survey of some new ceramic materials, such as the magnetic ceramics "Ferroxcube" and "Ferroxdure", the systems $TiO_2$-$Al_2O_3$ and $NaCl_1$-$CaCl_2$ which have interesting dielectric properties, and semiconductors such as $Fe_2O_3$ - $TiO_2$.

**2145:** J. Smit and J. Volger: Spontaneous Hall effect in ferromagnetics (Phys. Rev. **92**, 1576-1577, 1953, No. 6).

The coefficients $A_H$ (normal Hall coefficient) and $\varrho_{SH}$ occurring in the equation of the Hall effect:

$$E_y/i_x = A_H \cdot B_z + \varrho_{SH}$$

have been measured for 16 examples of Ni and Ni alloys in fields up to 14 000 gauss (1.4 Wb/m$^2$) at three temperatures (20 °K, 77 °K, 290 °K). The "spontaneous Hall resistivity" $\varrho_{SH}$ proves to be roughly proportional to the ordinary resistivity $\varrho$ and smaller by a factor of the order ($10^{-2}$) and vanishes for the purest Ni sample at T = 20 °K. It is shown that this last fact is in accordance with the theory.

**2146:** H. P. J. Wijn: Frequency-dependence of magnetization processes in ferrites and its relation to the distortion caused by ferrite cores (Soft magnetic materials for telecommunications, Pergamon Press, London 1953, pp. 51-63).

Magnetization curves of ferrites have been measured as a function of frequency and it appears that two dispersion mechanisms occur. The relation between the dispersion frequency of the initial permeability $\mu_i$ and the value of $\mu_i$ at low frequency has been investigated for samples of nickel-zinc ferrite with increasing zinc-content; the greater part of $\mu_i$ is caused by a rotation of the spins in the Weiss domains. The dispersion of the permeability corresponding to magnetic fields of the order of magnitude of the coercive force of the ferrite always occurs at a lower frequency, and is attributed to a relaxation of the irreversible domain-wall displacements. As a consequence of the last-mentioned dispersion, the distortion caused by coils with a ferrite core can decrease with frequency, and even vanish. Total losses of the ferrite are given as a function of induction and frequency, and the relation between distortion and hysteresis resistance is discussed.

**2147:** J. D. Fast and M. B. Verrijp: Diffusion of nitrogen in iron (J. Iron and Steel Inst. **176**, 24-27, 1954, No. 1).

Diffusion coefficients of nitrogen in $\alpha$-iron at 500° and 600° C are derived from the desorption rates of nitrogen from iron wires in hydrogen, using internal friction measurements to determine concentration ratios. Extrapolation to 21.5° and 9.5° gives values of $D$ that correspond well with those calculated from internal friction measurements at these temperatures and thus strongly support the interpretation by Snoek and Polder of the damping caused by carbon and nitrogen. The combined measurements yield for the diffusivity of nitrogen in $\alpha$-iron:

$$D = 6.6 \times 10^{-3} \exp(-18\,600/RT) \text{ cm}^2/\text{s}.$$

Determination of the diffusion coefficient of nitrogen in $\gamma$-iron at 950 °C shows that diffusion in $\gamma$-iron is much slower than in $\alpha$-iron.

**2148*:** H. Bruining: Physics and applications of secondary emission (Pergamon Press, London 1954, xii + 178 pp., 129 figs.)

In this book a survey of the physics and applications of secondary electron emission is given. Physical aspects are discussed in the first seven chapters. Ch. 1: Introduction; Ch. 2: Methods and measurements; Ch. 3 & 4: Review of results (metals and compounds); Ch. 5: Influence of externally adsorbed foreign atoms and ions; Ch. 6 and 7: Mechanism of secondary electron emission. A complete theory does not yet exist; a survey is given of various approaches to the problem. The final three chapters deal with applications. Ch. 8: Electron multiplication; Ch. 9. Elimination of disturbing effects caused by secondary emission; Ch. 10: Storage devices (for storage of information in the form of electric charges on an insulating surface, utilizing secondary emission). A list containing about 400 references is included.

**2149:** K. ter Haar and J. Bazen: The titration of Al with "Complexone III" at $p$H 3.5 (Anal. chim. Acta **10**, 23-28, 1954, No. 1).

A titration procedure for Al, based on the reaction of Al and "Complexone III" (disodium salt of ethylenediamine-tetra-acetic acid, ("versene") at $p$H 3.5-4.3 is described. After adding an excess of "Complexone III", the excess is back-titrated with thorium nitrate, using "Alizarin S" as an indicator. The reaction is nor exactly stoichiometric, but nevertheless quite reproducible; the correction factor is about 1%.